Lecture 1 (Sep. 6, 2017)

1.1 A "Weird" Example in Quantum Mechanics

Suppose we have an atom with spin-$\frac{1}{2}$, and we measure the z-component of its spin, S^z. The first claim of quantum mechanics that differs from classical mechanics is that we will always find one of the two values, $\pm\frac{\hbar}{2}$, where $\hbar = \frac{h}{2\pi}$.

The first thing we should wonder is, "How can we reasonably make this measurement?" There is a standard setup for such a measurement known as the Stern–Gerloch filter. We shoot a beam of these atoms through a region of spatially-varying magnetic field $\boldsymbol{B} = B(z)\hat{\boldsymbol{z}}$, such as the region between a square north-pole magnet and a pointed south-pole magnet. When travelling through this region, spin-up and spin-down particles will deflect in opposite directions, allowing us to filter the particles by their spin.

The reason this works is that the energy of the magnetic dipole moment of the atom in the magnetic field is

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\mu_z B_z \,, \tag{1.1}$$

and the force on the atom in the z-direction is thus

$$F_z = -\frac{\partial E}{\partial z} = \mu_z \frac{\mathrm{d}B_z}{\mathrm{d}z} \,. \tag{1.2}$$

Thus, states with different values of μ_z will deflect in different directions along the z-axis.

Now, let's imagine we have such an apparatus that can filter out the components of spin. We could modify this apparatus by changing its axis of orientation, and in this way we could create filters for spin along each of the x-, y-, and z-axes. Imagine then a sequence of filtering protocols to which we can subject our beam of atoms.

First, suppose we use this device to select out the $S^z = +\frac{\hbar}{2}$ atoms, and then measure S^z once again. This can be done by feeding the spin-up output of one z-oriented Stern–Gerloch filter into the input of a second z-oriented Stern–Gerloch filter. With this setup, we will find that 100% of the output of the second Stern–Gerloch filter will be spin-up.

Now consider altering this setup so that the second filter is oriented along the x-axis, so we are sending the atoms through a z-axis filter and then an x-axis filter. The claim in quantum mechanics is that we will now find that $S^x = +\frac{\hbar}{2}$ 50% of the time, and $S^x = -\frac{\hbar}{2}$ 50% of the time.

Now let's complicate matters further by taking the $S^x = +\frac{\hbar}{2}$ output of the x-axis filter and feeding it into yet another z-axis filter. The claim is that we will find $S^z = +\frac{\hbar}{2}$ 50% of the time, and $S^z = -\frac{\hbar}{2}$ 50% of the time. We have filtered out atoms that are only spin-up along the z-axis, then from among these we have filtered out only those that are spin-up along the x-axis, and then measured the spin along the z-axis again, and found that the atoms are equally likely to be spin-up and spin-down along the z-axis.

The reason for this phenomenon is that in quantum mechanics, S^z and S^x are *incompatible observables*, and cannot be measured simultaneously with complete accuracy. As far as we know, this behavior is inconsistent with a classical description of the state of a spin-$\frac{1}{2}$ particle.

1.2 The Fundamental Postulates of Quantum Mechanics

We will take the following as postulates of a quantum mechanical system:

1. The state of a quantum mechanical system at a fixed time t is given by a mathematical object known as a *vector* (or more precisely, a *ray*) $|\psi\rangle$, which is an element of a *complex Hilbert space* $\mathcal{H}$.

2. Observers cannot measure the state directly, but rather can only measure *observables*, which are *Hermitian operators* that act in the Hilbert space $\mathcal{H}$, and whose eigenvectors form a complete set.

3. A measurement of an observable $\hat{A}$ returns one of its eigenvalues.

   a. If $\hat{A}$ is measured in a state $|\psi\rangle$, then the probability that $A = a$ is given by

   $$\langle\psi|M_a|\psi\rangle\,, \tag{1.3}$$

   where

   $$M_a = \sum_{j:a_j=a} |a_j\rangle\langle a_j| \tag{1.4}$$

   is the *measurement operator*. Here, the sum is over all states $|a_j\rangle$ with eigenvalue a, i.e., $\hat{A}|a_j\rangle = a|a_j\rangle$.

   b. After measurement, the system is in the state

   $$|\tilde{\psi}\rangle \propto M_a|\psi\rangle\,. \tag{1.5}$$

   This is an often-discussed feature of quantum mechanics, referred to as the "collapse" of the wavefunction.

4. Time evolution is carried out by a map

   $$|\psi(t)\rangle \mapsto |\psi\big(t'\big)\rangle = U(t',t)|\psi(t)\rangle\,, \tag{1.6}$$

   where $U(t',t)$ is the *time-evolution operator*, which is a *unitary operator*, $U^\dagger U = \mathbb{1}$. The infinitesimal form of the time-evolution operator, when t' and t are infinitesimally separated, takes the form

   $$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle\,, \tag{1.7}$$

   where $\hat{H}$ is a Hermitian operator called the *Hamiltonian*.

1.3 Mathematical Preliminaries

1.3.1 Hilbert Spaces

We will begin by recalling the definition of a vector space. A *vector space* V is a collection of objects $\{|\alpha\rangle\}$ with several properties:

1. There is a well-defined *addition* operator '+' defined on this collection, so that for any $|\alpha\rangle, |\beta\rangle \in V$, there exists a unique state

   $$|\alpha\rangle + |\beta\rangle := |\gamma\rangle \in V\,. \tag{1.8}$$

   a. We demand that the addition operator is *commutative*: for all $|\alpha\rangle, |\beta\rangle \in V$,

   $$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle\,. \tag{1.9}$$

   b. We demand that the addition operator is *associative*: for all $|\alpha\rangle, |\beta\rangle, |\gamma\rangle \in V$,

   $$(|\alpha\rangle + |\beta\rangle) + |\gamma\rangle = |\alpha\rangle + (|\beta\rangle + |\gamma\rangle)\,. \tag{1.10}$$

2. We require the existence of a *null vector* $|0\rangle \in V$ with the property that

$$|0\rangle + |\alpha\rangle = |\alpha\rangle \tag{1.11}$$

for all $|\alpha\rangle \in V$.

3. For all $|\alpha\rangle \in V$, we require the existence of an *inverse* state $-|\alpha\rangle \in V$ with the property

$$|\alpha\rangle + (-|\alpha\rangle) = |0\rangle\,. \tag{1.12}$$

4. There is a well-defined *scalar multiplication* operation: for some field $\mathbb{F}$ (such as $\mathbb{R}$ or $\mathbb{C}$), we can multiply any state $|\alpha\rangle \in V$ by any *scalar* $c \in \mathbb{F}$, yielding a unique state $c|\alpha\rangle \in V$.

   a. For all $c, d \in \mathbb{F}$ and $|\alpha\rangle \in V$, scalar multiplication must satisfy

$$c(d|\alpha\rangle) = (cd)|\alpha\rangle\,, \tag{1.13}$$

   where cd indicates multiplication in the field $\mathbb{F}$.

   b. For all $|\alpha\rangle \in V$, the multiplicative identity element $1 \in \mathbb{F}$ must satisfy

$$1|\alpha\rangle = |\alpha\rangle\,. \tag{1.14}$$

   c. Scalar multiplication must distribute over addition in the vector space: for all $c \in \mathbb{F}$ and $|\alpha\rangle, |\beta\rangle \in V$, we must have

$$c(|\alpha\rangle + |\beta\rangle) = c|\alpha\rangle + c|\beta\rangle \tag{1.15}$$

   d. Scalar multiplication must distribute over addition in the field of scalars: for all $c, d \in F$ and $|\alpha\rangle \in V$, we must have

$$(c + d)|\alpha\rangle = c|\alpha\rangle + d|\alpha\rangle\,. \tag{1.16}$$

If the field of scalars is $\mathbb{F}$, then we refer to V as a "vector space over $\mathbb{F}$." A vector space V is called a *real vector space* if its field of scalars is the field of real numbers, $\mathbb{F} = \mathbb{R}$, and a *complex vector space* if its field of scalars is the field of complex numbers, $\mathbb{F} = \mathbb{C}$.

As a familiar example, we might consider the vectors in $\mathbb{R}^n$ of the form

$$\boldsymbol{v} = (v_1, \ldots, v_n)\,. \tag{1.17}$$

These form a real vector space under addition. We could also consider the space of states of a spin-$\frac{1}{2}$ particle,

$$c_+|+\rangle + c_-|-\rangle\,, \tag{1.18}$$

with $c_\pm \in \mathbb{C}$, which form a complex vector space. A more interesting example of a vector space is the space of all real functions $f(x)$ on the interval $[0, 1]$.

A subset $V \subset W$ of a vector space W is a *subspace* of W if it itself is a vector space over the field of scalars of W. A *ray* is a subspace of a vector space V of the form $\{c|\alpha\rangle \mid c \in \mathbb{F}\}$, for some $|\alpha\rangle \in V$.

A collection of vectors $|\alpha_1\rangle, \ldots, |\alpha_n\rangle \in V$ is *linearly independent* if and only if the statement

$$\sum_{j=1}^{n} c_j|\alpha_j\rangle = 0 \tag{1.19}$$

implies that $c_j = 0$ for all $j = 1, \ldots, n$. A collection of vectors $\{|\alpha_j\rangle\}$ is *maximally linearly independent* if it is linearly independent and there exists no state $|\beta\rangle \in V$ such that the collection $\{|\alpha_j\rangle, |\beta\rangle\}$ is linearly independent. In this case, $\{|\alpha_j\rangle\}$ is called a *basis* for V, and the number of elements n in the collection is called the *dimension* of V, denoted $\dim V$. The dimension of a vector space can be finite or infinite, and if infinite, can be either countable (also called denumerable) or uncountable.

Let's consider some examples. A finite-dimensional vector space with which we are already familiar is $\mathbb{R}^n$, which has $\dim \mathbb{R}^n = n$. An example of a infinite-dimensional vector space is the space of square-integrable functions on the interval $[0, 1]$ with $f(0) = f(1) = 0$. This is a countable vector space: we can expand such a function $f(x)$ as

$$f(x) = \sum_{k=1}^{\infty} c_k \sin(\pi k x)\,, \tag{1.20}$$

which provides a countable basis of functions for this space.

Similarly, we can consider the vector space of square integrable functions on $(-\infty, \infty)$. More surprisingly, this vector space also has countable dimension. A clean way to see this is to produce a countable basis for this space: one such basis is the set of eigenfunctions of the simple harmonic oscillator. These eigenfunctions are all square-integrable, and any other wavefunction can be expanded in this basis. The oscillator eigenfunctions are labelled by integers, so this shows that this space is countable. This may seem confusing, because another basis we might use is the position basis, where our basis elements are Dirac delta functions at each position; the problem with this approach is that the Dirac delta function is not itself square-integrable.

Lecture 2 (Sep. 11, 2017)

2.1 More Relevant Math

2.1.1 Inner Products

Last time, we discussed the concept of a maximally linearly independent set, which is a set $\{|\alpha_j\rangle\}$ of vectors that are linearly independent, and such that there exists no $|\beta\rangle$ such that $\{|\alpha_j\rangle, |\beta\rangle\}$ is linearly independent. In this case, $\{|\alpha_j\rangle\}$ forms a basis for V. If $|\alpha_1\rangle, \ldots, |\alpha_n\rangle$ form a basis for V, then any vector $|\beta\rangle$ can be decomposed in terms of the basis vectors as

$$|\beta\rangle = \sum_i c_i |\alpha_i\rangle \,. \tag{2.1}$$

Next, we discuss the concept of an inner product (we will assume that V is a vector space $\mathbb{C}$). A (Hermitian) *inner product* is a map $(\cdot,\cdot)\colon V \times V \to \mathbb{C}$ with the following properties:

1. For all $c \in \mathbb{C}$ and $|\alpha\rangle, |\beta\rangle \in V$, we have

$$\begin{aligned} (|\alpha\rangle, c|\beta\rangle) &= c(|\alpha\rangle, |\beta\rangle) \,, \\ (c|\alpha\rangle, |\beta\rangle) &= c^*(|\alpha\rangle, |\beta\rangle) \,, \end{aligned} \tag{2.2}$$

   where c^* is the complex conjugate of c.

2. For all $|\alpha\rangle, |\beta\rangle \in V$,

$$(|\alpha\rangle, |\beta\rangle)^* = (|\beta\rangle, |\alpha\rangle) \,. \tag{2.3}$$

3. For all $|\alpha\rangle, |\beta\rangle, |\beta'\rangle \in V$,

$$(|\alpha\rangle, |\beta\rangle + |\beta'\rangle) = (|\alpha\rangle, |\beta\rangle) + (|\alpha\rangle, |\beta'\rangle) \,. \tag{2.4}$$

4. For all $|\alpha\rangle \in V$,

$$(|\alpha\rangle, |\alpha\rangle) \geq 0 \,, \tag{2.5}$$

   and if $(|\alpha\rangle, |\alpha\rangle) = 0$, then $|\alpha\rangle$ is the zero ket, $|\alpha\rangle = 0$.

Let's consider some examples. First, consider $V = \mathbb{C}^n$, whose vectors are of the form

$$|z\rangle = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \quad z_1, \ldots, z_n \in \mathbb{C} \,. \tag{2.6}$$

One definition of an inner product on this space is

$$(|z\rangle, |w\rangle) = z_1^* w_1 + z_2^* w_2 + \cdots + z_n^* w_n \,, \tag{2.7}$$

which we can check satisfies all of the required properties of an inner product.

A more exciting example is the vector space of square-integrable functions on the interval $[0, 1]$,

$$V = \left\{ f\colon [0,1] \to \mathbb{C} \;\middle|\; \int_0^1 \mathrm{d}x \; f^*(x) f(x) < \infty \right\}, \tag{2.8}$$

as we discussed in the last lecture, which can be endowed with the inner product

$$(f, g) = \int_0^1 \mathrm{d}x \; f^*(x) g(x) \,. \tag{2.9}$$

Now we introduce a bit of terminology. The vectors we have been discussing are referred to as *kets*. Two kets are *orthogonal* if $(|\alpha\rangle, |\beta\rangle) = 0$. The *norm* of a ket is given by

$$N := \sqrt{(|\alpha\rangle, |\alpha\rangle)} := \||\alpha\rangle\| \,. \tag{2.10}$$

Using this norm, we can normalize any nonzero ket by defining

$$|\tilde{\alpha}\rangle = \frac{1}{N} |\alpha\rangle \,, \tag{2.11}$$

which has

$$(|\tilde{\alpha}\rangle, |\tilde{\alpha}\rangle) = 1 \,. \tag{2.12}$$

2.1.2 Dual Space

Now we introduce the dual space. The space V that we have described so far is the space of kets $|\alpha\rangle$. The *dual space* V^* is the space of objects called *bras*, which is "dual" to the ket space. To every $|\alpha\rangle \in V$, we associate a bra $\langle\alpha| \in V^*$, and define an operation between bras and kets,

$$\langle\alpha|\beta\rangle := (|\alpha\rangle, |\beta\rangle) \,. \tag{2.13}$$

This shows us the motivation for these names: when we combine a bra and a ket, we get a "bra(c)ket."

Consider the space of linear functionals $\gamma \colon V \to \mathbb{C}$. We introduce the notation $\gamma \to \langle\gamma|$, such that for any basis $\{|\alpha_j\rangle\}$, the statement

$$\gamma \colon |\alpha_j\rangle \mapsto c_j \tag{2.14}$$

implies that

$$\langle\gamma|\alpha_j\rangle = c_j \,. \tag{2.15}$$

We see that the dual space is just the space of linear functionals on V. The $\{\gamma\}$ form an n-dimensional vector space V^*, where $n = \dim V$. We then have a duality $V \leftrightarrow V^*$, given by $|\alpha\rangle \leftrightarrow \langle\alpha|$. Note that scalars are conjugated under this duality, so that for all $c \in \mathbb{C}$ and $|\alpha\rangle \in V$, $c|\alpha\rangle \leftrightarrow c^*\langle\alpha|$.

2.1.3 Orthonormal Bases

An *orthonormal basis* is a basis $\{|\phi_i\rangle\}$ such that

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} \,, \tag{2.16}$$

where δ_{ij} is the Kronecker delta. We know that we can write any ket in the form

$$|\alpha\rangle = \sum_i c_i |\phi_i\rangle \,. \tag{2.17}$$

If the $|\phi_i\rangle$ form an orthonormal basis, then we find that $c_i = \langle\phi_i|\alpha\rangle$. Thus, if we have an orthonormal basis, then we can write any ket in the form

$$|\alpha\rangle = \sum_i |\phi_i\rangle\langle\phi_i|\alpha\rangle. \tag{2.18}$$

These is an example of a *completeness relation*.

2.1.4 Operators

Recall the second postulate of quantum mechanics: observables are represented in quantum mechanics as Hermitian operators acting on the Hilbert space $\mathcal{H}$ (from here on out, we will consider states in the Hilbert space $\mathcal{H}$ rather than an arbitrary vector space V). In order to understand this statement, we first need to understand what an operator is, and then we must understand what a Hermitian operator is.

An *operator* is an object that acts on a ket and returns another ket. Thus, if X is an operator, then X acts on a ket $|\alpha\rangle \in \mathcal{H}$ as

$$X(|\alpha\rangle) := X|\alpha\rangle \in \mathcal{H}\,, \tag{2.19}$$

where the right-hand side is a naming convention. Two operators A and B are said to be *equal* ($A = B$) if and only if $A|\alpha\rangle = B|\alpha\rangle$ for all $|\alpha\rangle \in \mathcal{H}$.

We introduce an addition operation '+' on the space of operators, which is commutative and associative: for all operators X, Y, Z, we have

$$\begin{aligned} X + Y &= Y + X\,, \\ X + (Y + Z) &= (X + Y) + Z\,. \end{aligned} \tag{2.20}$$

We will primarily be interested in linear operators: a *linear* operator X satisfies the property

$$X(c_\alpha|\alpha\rangle + c_\beta|\beta\rangle) = c_\alpha X|\alpha\rangle + c_\beta X|\beta\rangle \tag{2.21}$$

for all $c_\alpha, c_\beta \in \mathbb{F}$ and $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$. We can also define the notion of an anti-linear operator: an *anti-linear* operator X satisfies

$$X(c_\alpha|\alpha\rangle + c_\beta|\beta\rangle) = c_\alpha^* X|\alpha\rangle + c_\beta^* X|\beta\rangle \tag{2.22}$$

for all $c_\alpha, c_\beta \in \mathbb{F}$ and $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$.

We can also define an action of these same operators on the dual space, which will be a right action of the form $\langle\beta|X$ for $\langle\beta| \in \mathcal{H}^*$. We define this object via its inner product with kets, by requiring that it satisfy the property

$$(\langle\beta|X)|\alpha\rangle = \langle\beta|(X|\alpha\rangle) \tag{2.23}$$

for all $|\alpha\rangle \in \mathcal{H}$. Using an orthonormal basis $\{|\phi_j\rangle\}$, we can then express this bra as

$$\langle\beta|X = \sum_j \langle\beta|X|\phi_j\rangle\langle\phi_j|\,. \tag{2.24}$$

We also define multiplication of operators, such that for all operators X, Y and $|\alpha\rangle \in \mathcal{H}$,

$$(XY)|\alpha\rangle = X(Y|\alpha\rangle)\,. \tag{2.25}$$

This multiplication is not necessarily commutative, $XY \neq YX$, but it is associative,

$$X(YZ) = (XY)Z = XYZ\,. \tag{2.26}$$

2.1.5 Operators as Matrices in a Given Basis

In a given basis, the action of an operator can be expressed by a matrix. To see this, we first define the identity operator $\mathbb{1}$, which satisfies

$$\mathbb{1}|\alpha\rangle = |\alpha\rangle \tag{2.27}$$

for all $|\alpha\rangle \in \mathcal{H}$. Given an orthonormal basis $|\{a'\}\rangle$ (notation from Sakurai), the identity operator can be expressed in the form

$$\mathbb{1} = \sum_{a'} |a'\rangle\langle a'| \,. \tag{2.28}$$

We can use Eq. (2.28) to write any operator X in the form

$$\begin{aligned} X &= \mathbb{1}X\mathbb{1} \\ &= \left(\sum_{a'} |a'\rangle\langle a'|\right) X \left(\sum_{a''} |*a''\rangle\langle a''|\right) \\ &= \sum_{a',a''} |a'\rangle\langle a'|X|a''\rangle\langle a''| \,. \end{aligned} \tag{2.29}$$

Note that $\langle a'|X|a''\rangle \in \mathbb{C}$ is simply a number. Thus, all information about X is equivalently contained in the complex numbers $\langle a'|X|a''\rangle$, given any orthonormal basis $|\{a'\}\rangle$. This defines an $n \times n$ matrix with complex entries corresponding to each operator X. For this reason, we will often use the words "operator" and "matrix" interchangeably if the chosen basis is clear, even though the concept of an operator is more fundamental.

2.1.6 Adjoint Operators

For each operator X, we define its *adjoint operator* (also called the *Hermitian conjugate*) $X^\dagger$ to be the operator satisfying

$$X^\dagger|\alpha\rangle \leftrightarrow \langle\alpha|X \tag{2.30}$$

under the duality $\mathcal{H} \leftrightarrow \mathcal{H}^*$. If $X^\dagger = X$, then the operator X is called *self-adjoint* or *Hermitian*. (Note that in infinite-dimensional spaces, the concepts of self-adjointness and Hermiticity are technically distinct. We may discuss this later in the course.) Note that $(XY)^\dagger = Y^\dagger X^\dagger$. The proof of this fact is straightforward: For arbitrary $|\alpha\rangle \in \mathcal{H}$, consider the kets defined by

$$|\beta\rangle = Y|\alpha\rangle \,, \quad |\gamma\rangle = X|\beta\rangle \,. \tag{2.31}$$

Their dual bras are then defined by

$$\langle\alpha|Y^\dagger = \langle\beta| \,, \quad \langle\beta|X^\dagger = \langle\gamma| \,. \tag{2.32}$$

From these definitions, we have

$$XY|\alpha\rangle = |\gamma\rangle \,, \quad \langle\alpha|Y^\dagger X^\dagger = \langle\gamma| \,. \tag{2.33}$$

Taking the dual of the first equation in (2.33) yields

$$\langle\alpha|(XY)^\dagger = \langle\gamma| \,, \tag{2.34}$$

which we can compare with the second equation in (2.33) to yield $(XY)^\dagger = Y^\dagger X^\dagger$.

This same concept has a well-defined meaning for matrices. As we have seen, any operator has the same information content as an $n \times n$ matrix, where $n = \dim \mathcal{H}$. If M is a matrix corresponding to the operator X, then $X^\dagger$ corresponds to the matrix found by conjugating the entries of the transpose M^{T}. This matrix is denoted as $M^\dagger$, and is also called the *Hermitian conjugate*. A matrix satisfying $M^\dagger = M$ is called a *Hermitian* matrix.

2.1.7 Operator Examples

As a first example of an operator, we define the outer product. Given a ket $|\alpha\rangle \in \mathcal{H}$ and a bra $\langle\beta| \in \mathcal{H}^*$, we define the operator

$$\mathcal{O} = |\alpha\rangle\langle\beta| \tag{2.35}$$

to be their *outer product*. This is an operator: for all $|\gamma\rangle \in \mathcal{H}$, we have

$$\mathcal{O}|\gamma\rangle = |\alpha\rangle\langle\beta|\gamma\rangle\,. \tag{2.36}$$

Taking the dual of this equation, we find

$$\begin{aligned}\langle\gamma|\mathcal{O}^\dagger &= \langle\beta|\gamma\rangle^*\langle\alpha| \\ &= \langle\gamma|\beta\rangle\langle\alpha|\,,\end{aligned} \tag{2.37}$$

and so we conclude that

$$\mathcal{O}^\dagger = |\beta\rangle\langle\alpha|\,. \tag{2.38}$$

As a second example, consider the Hilbert space $\mathcal{H}$ of square-integrable functions on $(-\infty, \infty)$, with inner product

$$\langle f_1|f_2\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\; f_1^*(x) f_2(x) \tag{2.39}$$

for $|f_1\rangle, |f_2\rangle \in \mathcal{H}$. Now consider the operator $A = \frac{\mathrm{d}}{\mathrm{d}x}$. Using integration by parts, we see that

$$\int_{-\infty}^{\infty} \mathrm{d}x\; f_1^*(x)\frac{\mathrm{d}}{\mathrm{d}x} f_2(x) = \int_{-\infty}^{\infty} \mathrm{d}x \left(-\frac{\mathrm{d}}{\mathrm{d}x} f_1^*(x)\right) f_2(x)\,, \tag{2.40}$$

telling us that

$$\left\langle f_1 \middle| \frac{\mathrm{d}}{\mathrm{d}x} \middle| f_2 \right\rangle = -\left\langle f_2 \middle| \frac{\mathrm{d}}{\mathrm{d}x} \middle| f_1 \right\rangle^*. \tag{2.41}$$

Thus,

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^\dagger = -\frac{\mathrm{d}}{\mathrm{d}x}\,. \tag{2.42}$$

Operators satisfying $A^\dagger = -A$ are called *anti-Hermitian*. We can make this into a Hermitian operator by multiplying it by the imaginary unit (or its negative):

$$\left(-i\frac{\mathrm{d}}{\mathrm{d}x}\right)^\dagger = -i\frac{\mathrm{d}}{\mathrm{d}x}\,. \tag{2.43}$$

You may recognize this as the momentum operator in one-dimensional quantum mechanics.

A third example is the identity operator $\mathbb{1}$, which we have already seen.

For some operators, we can define an inverse operator: for an operator A, its inverse operator A^{-1} (if it exists) satisfies

$$A^{-1}A = AA^{-1} = \mathbb{1}\,. \tag{2.44}$$

Inverse operators are not guaranteed to exist.

Of special interest to us are unitary operators: a *unitary* operator U is one that satisfies $U^{-1} = U^\dagger$. These are useful because they preserve the inner product:

$$\left(\langle\beta|U^\dagger\right)(U|\alpha\rangle) = \langle\beta|U^{-1}U|\alpha\rangle = \langle\beta|\alpha\rangle\,. \tag{2.45}$$

We are also interested in *projection* operators, which are operators A satisfying $A^2 = A$. An example of a projection operator is $A = |\alpha\rangle\langle\alpha|$, for some $|\alpha\rangle \in \mathcal{H}$.

2.1.8 Eigenstates and Eigenvalues

If, for some operator A and ket $|\alpha\rangle \in \mathcal{H}$, we have

$$A|\alpha\rangle = a|\alpha\rangle\,, \tag{2.46}$$

with $a \in \mathbb{C}$, then $|\alpha\rangle$ is an *eigenket* or *eigenstate* of the operator A, with *eigenvalue* a. The *spectrum* of an operator A is the set of its eigenvalues $\{a\}$.

We now prove an important theorem regarding the eigenkets and eigenvalues of Hermitian operators.

Theorem 1 (Spectral Theorem)**.** *If A is a Hermitian operator, $A = A^\dagger$, then*

1. *all eigenvalues a_i of A are real, and*
2. *eigenkets of A with distinct eigenvalues are orthogonal.*

Proof. Suppose that for some Hermitian operator A and kets $|a'\rangle, |a''\rangle \in \mathcal{H}$ we have

$$A|a'\rangle = a'|a'\rangle\,, \tag{2.47a}$$
$$\langle a''|A = \langle a''|(a'')^*\,. \tag{2.47b}$$

Multiplying Eq. (2.47a) by $\langle a''|$ on the left and multiplying Eq. (2.47b) by $|a'\rangle$ on the right yields

$$\langle a''|A|a'\rangle = a'\langle a''|a'\rangle\,, \tag{2.48a}$$
$$\langle a''|A|a'\rangle = (a'')^*\langle a''|a'\rangle\,. \tag{2.48b}$$

Comparing Eqs. (2.48a) and (2.48b), we see that

$$\big[(a'')^* - a'\big]\langle a''|a'\rangle = 0\,. \tag{2.49}$$

If $|a'\rangle = |a''\rangle$, then this proves that $(a')^* = a'$, which means that $a' \in \mathbb{R}$. If $|a'\rangle \neq |a''\rangle$ and we have $a' \neq a''$, then this proves that $\langle a''|a'\rangle = 0$, i.e., $|a'\rangle$ and $|a''\rangle$ are orthogonal. This completes the proof. □

If we normalize each eigenket $|a_i\rangle$ of a Hermitian operator, so that $\langle a_i|a_i\rangle = 1$, and have chosen the eigenkets so that distinct eigenkets with the same eigenvalue are orthogonal, then Thm. 1 tells us that we have

$$\langle a_i|a_j\rangle = \delta_{ij}\,, \tag{2.50}$$

i.e., the normalized eigenkets are orthonormal. We can then decompose

$$A = \sum_a a|a\rangle\langle a|\,. \tag{2.51}$$

We can check that for some eigenket $|b\rangle$, we have

$$A|b\rangle = \sum_a a|a\rangle\langle a|b\rangle = \sum_a a|a\rangle\delta_{ab} = b|b\rangle\,. \tag{2.52}$$

Note that the equation

$$\mathbb{1} = \sum_a |a\rangle\langle a| \tag{2.53}$$

is simply a special case of Eq. (2.51).

Lecture 3 (Sep. 13, 2017)

3.1 Even More Math

3.1.1 More on Matrix Representations

Last time, we described that in a given basis, there is an exact correspondence between $n \times n$ matrices and operators, where n is the dimension of the Hilbert space. Let $\{|a\rangle\}$ form an orthonormal basis, so that any state $|\alpha\rangle$ can be expanded as

$$|\alpha\rangle = \sum_a \langle a|\alpha\rangle |a\rangle := \sum_a c_{\alpha a}|a\rangle \,. \tag{3.1}$$

Similarly, any operator X can be expressed in the form

$$X = \sum_{a_i, a_j} |a_i\rangle\langle a_i|X|a_j\rangle\langle a_j| \,. \tag{3.2}$$

In this basis, all of the information of the operator X is contained in the matrix

$$X_{ij} = \langle a_i|X|a_j\rangle \,. \tag{3.3}$$

The terminology is that the X_{ij} are *matrix elements* of X between states $|a_i\rangle$ and $|a_j\rangle$. As an exercise, you can check that the matrix elements of the product operator XY are given by $\sum_j X_{ij}Y_{jk}$, where X_{ij} and Y_{jk} are the matrix elements of the operators X and Y in some basis, respectively.

We define the *trace* of an operator A as

$$\mathrm{Tr}\, A = \sum_i \langle a_i|A|a_i\rangle = \sum_i A_{ii} \,. \tag{3.4}$$

If the $|a_i\rangle$ are chosen to be eigenvectors of A, then $A_{ii} = a_i$, so the trace becomes

$$\mathrm{Tr}\, A = \sum_i a_i \,. \tag{3.5}$$

This is a statement of the familiar fact that the trace of an operator is the sum of its eigenvalues.

3.1.2 Unitary Transformations

Suppose we are given two orthonormal bases $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$. How are these bases related? We can define an operator U by the action

$$U|a_i\rangle = |b_i\rangle \,. \tag{3.6}$$

This implies that

$$\langle b_i| = \langle a_i|U^\dagger \,, \tag{3.7}$$

by definition of the adjoint.

Note that we can write

$$U = U\mathbb{1} = U\sum_i |a_i\rangle\langle a_i| = \sum_i |b_i\rangle\langle a_i| \,, \tag{3.8}$$

so the operator U is simply the sum of the outer products of corresponding vectors from each basis. Then we have

$$U^\dagger = \sum_i |a_i\rangle\langle b_i| \,, \tag{3.9}$$

from which we find that

$$\begin{aligned} UU^\dagger &= \sum_{i,j} |b_i\rangle\langle a_i|a_j\rangle\langle b_j| \\ &= \sum_i |b_i\rangle\langle b_i| \\ &= \mathbb{1} \,. \end{aligned} \tag{3.10}$$

Thus, we see that U is unitary. Unitary transformations are precisely those that transform from one orthonormal basis to another.

Consider now a vector $|\alpha\rangle$ and two distinct bases $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$. We can express $|\alpha\rangle$ in two ways as

$$|\alpha\rangle = \sum_i c_i |a_i\rangle = \sum_i d_i |b_i\rangle \,. \tag{3.11}$$

How are these two sets of coefficients $\{c_i\}$ and $\{d_i\}$ related? Using the definition of U, we can write

$$\begin{aligned} |\alpha\rangle &= \sum_j d_j |b_j\rangle \\ &= \sum_j d_j U|a_j\rangle \\ &= \sum_{i,j} d_j |a_i\rangle\langle a_i|U|a_j\rangle \,. \end{aligned} \tag{3.12}$$

Thus, we see that

$$c_{ij} = \sum_{i,j} \langle a_i|U|a_j\rangle d_j = U_{ij} d_j \,, \tag{3.13}$$

where we have introduced the shorthand notation in which repeated indices are assumed to be summed over (from now on, we will explicitly state the cases when we are not using this convention). Note that the U_{ij} are simply the matrix elements of the operator U.

A similar approach can be used to show that for any operator that can be expressed in two different bases as

$$X = \sum_i |a_i\rangle X_{ij} \langle a_j| = \sum_{k\ell} |b_k\rangle Y_{k\ell} \langle b_\ell| \,, \tag{3.14}$$

the matrix elements in the two different bases are related by

$$X_{ij} = U_{ik} Y_{k\ell} U^\dagger_{\ell j} \,. \tag{3.15}$$

We leave the proof as an exercise.

3.1.3 Diagonalization of Hermitian Operators

Theorem 2. *A Hermitian matrix $H_{ij} = \langle \phi_i|H|\phi_j\rangle$ can always be diagonalized by a unitary transformation.*

Proof. Consider a general orthonormal basis $\{|\phi_i\rangle\}$, and let $\{|h_i\rangle\}$ be the orthonormal basis of eigenstates of the operator H — such a basis exists because H is Hermitian. Because these are both orthonormal bases, there exists a unitary transformation U such that $|h_i\rangle = U|\phi_i\rangle$. Using this operator, we can then write

$$\delta_{ij} h_i = \langle h_i|H|h_j\rangle = \langle \phi_i|U^\dagger H U|\phi_j\rangle \,, \tag{3.16}$$

where the first expression on the left is *not* summed over i. Thus, we see that $U^\dagger_{ik} H_{k\ell} U_{\ell j}$ is a diagonal matrix. This completes the proof. □

3.1.4 Simultaneous Diagonalization

Theorem 3. *Two (diagonalizable) operators A, B are simultaneously diagonalizable if and only if $[A, B] = 0$, where $[\cdot, \cdot]$ is the commutator.*

Proof. Let there be a basis $\{|a_i\rangle\}$ for which the diagonalizable operators A, B have

$$A|a_i\rangle = a_i|a_i\rangle \,, \quad B|a_i\rangle = b_i|a_i\rangle \,. \tag{3.17}$$

This is a basis in which A and B are simultaneously diagonal. In this case, we see that

$$AB|a_i\rangle = a_i b_i|a_i\rangle = BA|a_i\rangle \,, \tag{3.18}$$

and so $AB = BA$. This proves the forward direction.

Now consider two diagonalizable operators A, B that commute, $AB = BA$, and let $\{|a_i\rangle\}$ be a basis of eigenvectors of A,

$$A|a_i\rangle = a_i|a_i\rangle \,. \tag{3.19}$$

Then we have

$$\begin{aligned} A(B|a_i\rangle) &= AB|a_i\rangle \\ &= BA|a_i\rangle \\ &= Ba_i|a_i\rangle \\ &= a_i(B|a_i\rangle) \,, \end{aligned} \tag{3.20}$$

so $B|a_i\rangle$ is an eigenket of A with eigenvalue a_i. In general, this means that B is block diagonal in the basis $\{|a_i\rangle\}$ (once we have ordered the basis so as to group eigenkets of A with the same eigenvalue), with each block corresponding to a single eigenvalue of A. We can then diagonalize B within each block, which will leave A diagonal, thereby simultaneously diagonalizing A and B. Visually, we have

$$A = \begin{pmatrix} a_1 & & & & & & & \\ & \ddots & & & & & & \\ & & a_1 & & & & & \\ & & & a_2 & & & & \\ & & & & \ddots & & & \\ & & & & & a_2 & & \\ & & & & & & \ddots & \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \end{pmatrix}, \tag{3.21}$$

where the B_i are $k_i \times k_i$ blocks, with k_i the number of occurrences of the eigenvalue a_i on the diagonal of A. Diagonalizing the block B_i only mixes eigenkets of A with the same eigenvalue, so we can diagonalize B while leaving A diagonal. This proves the backward direction. □

3.2 Measurement

Consider a quantum mechanical system. We know that the state of such a system consists of a normalized state $|\psi\rangle \in \mathcal{H}$, and observables are represented by Hermitian operators acting in the Hilbert space $\mathcal{H}$. Recall the third postulate, regarding measurement:

a. The possible results of measuring A are the eigenvalues of a_i of A.

b. Once a measurement is done, and the result is a_i for some i, the system "collapses" to an eigenket $|a_i\rangle$ with a sharp value of A: if we perform a measurement again instantaneously, we are guaranteed to get the same value as the first measurement.

c. The probability that a measurement gives the result $A = a_i$ is given by

$$\begin{aligned} \text{Prob}(A = a_i) &= \sum_{j:a_j=a_i} |\langle a_j|\psi\rangle|^2 \\ &= \sum_{j:a_j=a_i} \langle\psi|a_j\rangle\langle a_j|\psi\rangle \\ &= \langle\psi|M_{a_i}|\psi\rangle\,, \end{aligned} \tag{3.22}$$

where we have defined the *measurement operator*

$$M_{a_i} := \sum_{j:a_j=a_i} |a_j\rangle\langle a_j|\,. \tag{3.23}$$

This operator is the projector onto the subspace with $A = a_i$.

If we simplify to the case where there is only one eigenket with eigenvalue a_i. Then

$$\text{Prob}(A = a_i) = |\langle a_i|\psi\rangle|^2\,. \tag{3.24}$$

Historically, there has been a lot of worry about the collapse of the wavefunction. In the modern language, there is a way to understand this collapse in a very palatable way, which we may discuss later in the course.

The concept of measurement is worth pondering. Consider making a measurement on a state, and then immediately making another measurement on it. Can we actually do this? Say our measurement is to see if there is a photon hitting our detector. Prior to measurement, the state consists of a detector and a photon. After measurement, there is a click, and the photon is gone. How can we remeasure? We must separately consider measurements that destroy our state and ones that do not. Measurements that do not destroy the state are called "non-demolition measurements."

3.2.1 Comments

1. We have said that

$$\text{Prob}(A = a_i) = \sum_{j:a_j=a_i} |\langle a_j|\psi\rangle|^2\,. \tag{3.25}$$

This can only make sense if the value on the right-hand side is non-negative, and the sum of these probabilities over all eigenvalues is 1. It is clear that the right-hand side is non-negative,

as is the sum of non-negative numbers, and we see that

$$\begin{aligned}\sum_i \text{Prob}(A = a_i) &= \sum_j |\langle a_j|\psi\rangle|^2 \\ &= \sum_j \langle\psi|a_j\rangle\langle a_j|\psi\rangle \\ &= \langle\psi|\psi\rangle \\ &= 1\,,\end{aligned} \tag{3.26}$$

because the state $|\psi\rangle$ is assumed to be normalized.

2. For any observable A and state $|\psi\rangle$, the *expectation value* of A is

$$\begin{aligned}\langle A\rangle &:= \sum_{a_i} a_i \,\text{Prob}(A = a_i) \\ &= \sum_{a_i} a_i \sum_{j:a_j=a_i} \langle\psi|a_j\rangle\langle a_j|\psi\rangle \\ &= \langle\psi|A|\psi\rangle\,,\end{aligned} \tag{3.27}$$

where in the final line we have used

$$A = \sum_{a_i} a_i |a_i\rangle\langle a_i|\,. \tag{3.28}$$

3.3 Spin-$\frac{1}{2}$ Systems

Consider a spin-$\frac{1}{2}$ system. The state space is spanned by the eigenstates of, for example, S^z. We denote the state with $S^z = \frac{\hbar}{2}$ by $|+\rangle$ and the state with $S^z = -\frac{\hbar}{2}$ by $|-\rangle$. As a set, the Hilbert space is then

$$\mathcal{H} = \left\{ |\psi\rangle = c_+|+\rangle + c_i|-\rangle \;\middle|\; c_\pm \in \mathbb{C}, |c_+|^2 + |c_-|^2 = 1 \right\}. \tag{3.29}$$

This is a subspace of the two-dimensional complex vector space $\mathbb{C}^2$.

The states $|\psi\rangle$ and $e^{i\lambda}|\psi\rangle$, for some $\lambda \in \mathbb{R}$, are physically equivalent by definition (the Hilbert space is actually the space above quotiented by this equivalence relation), so the only physically relevant phase information in a state is the relative phase of the coefficients (c_+, c_-). We can then parameterize these coefficients in the form

$$c_+ = \cos\frac{\theta}{2}\,, \quad c_- = e^{i\phi}\sin\frac{\theta}{2}\,, \tag{3.30}$$

with $0 \le \theta < \pi$ and $0 \le \phi < 2\pi$. Specifying these two angles specifies the state exactly, and we note that specifying these angles is equivalent to specifying a point on the surface of the unit sphere S^2. This parametrization of the Hilbert space is known as the *Bloch sphere*. For example, the north pole of the Bloch sphere has $c_+ = 1, c_- = 0$, and so represents the state $|\psi\rangle = |+\rangle$, while the south pole has $c_+ = 0, c_- = 1$, and so represents the state $|\psi\rangle = |-\rangle$.

We will now define several important operators in this space. The identity operator is

$$\mathbb{1} = |+\rangle\langle+| + |-\rangle\langle-|\,, \tag{3.31}$$

which can be represented by a matrix

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.32}$$

Keep in mind that using an equals sign here is an abuse of notation, as the matrix representation of an operator is formally distinct from the operator itself. The operator S^z is given by

$$S^z = \frac{\hbar}{2}(|+\rangle\langle+| - |-\rangle\langle-|) = \frac{\hbar\sigma^z}{2} = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tag{3.33}$$

where σ^z is the third Pauli matrix. Similarly, we can write

$$\begin{aligned} S^x &= \frac{\hbar}{2}(|+\rangle\langle-| + |-\rangle\langle+|) = \frac{\hbar\sigma^x}{2} = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \\ S^y &= \frac{-i\hbar}{2}(|+\rangle\langle-| - |-\rangle\langle+|) = \frac{\hbar\sigma^y}{2} = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}. \end{aligned} \tag{3.34}$$

We can check explicitly (left as an exercise) that these operators satisfy

$$\begin{aligned} \left[S^a, S^b\right] &= i\hbar\epsilon^{abc}S^c \\ \left\{S^a, S^b\right\} &= \frac{\hbar^2}{2}\delta^{ab}\mathbb{1} \\ S^2 := S^aS^a &= \frac{3}{4}\hbar^2\mathbb{1} \\ \left[S^2, S^a\right] &= 0\,, \end{aligned} \tag{3.35}$$

where $[\cdot,\cdot]$ is the commutator, $\{\cdot,\cdot\}$ is the anticommutator, and ϵ^{abc} is the totally antisymmetric three-index tensor (known as the *Levi–Civita symbol*).

Lecture 4 (Sep. 18, 2017)

4.1 Measurement

4.1.1 Spin-$\frac{1}{2}$ Systems

Last time, we said that a general state in a spin-$\frac{1}{2}$ system can be written as

$$|\psi\rangle = c_+|+\rangle + c_-|-\rangle\,, \tag{4.1}$$

where $|+\rangle, |-\rangle \in \mathcal{H}$ are eigenstates of S^z and $c_+, c_- \in \mathbb{C}$. We noted that only the relative phase between c_+ and c_- was physically relevant.

Assume that $|\psi\rangle$ is normalized, so that $\langle\psi|\psi\rangle = 1$, which implies that

$$|c_+|^2 + |c_-|^2 = 1\,. \tag{4.2}$$

Now consider what happens when we measure S_z. We find

$$\begin{aligned} \text{Prob}\left(S^z = +\frac{\hbar}{2}\right) &= |\langle +|\psi\rangle|^2 = |c_+|^2\,, \\ \text{Prob}\left(S^z = -\frac{\hbar}{2}\right) &= |\langle -|\psi\rangle|^2 = |c_-|^2\,. \end{aligned} \tag{4.3}$$

We now want to consider a measurement of spin along an arbitrary axis $\hat{\boldsymbol{n}}$. In general, we have

$$\text{Prob}\left(\vec{S}\cdot\hat{\boldsymbol{n}} = \pm\frac{\hbar}{2}\right) = \left|\left\langle \vec{S}\cdot\hat{\boldsymbol{n}} = \pm\tfrac{\hbar}{2}\middle|\psi\right\rangle\right|^2, \tag{4.4}$$

where

$$\left|\vec{S}\cdot\hat{\boldsymbol{n}} = \pm\tfrac{\hbar}{2}\right\rangle \tag{4.5}$$

is the eigenket of the operator $\vec{S}\cdot\hat{\boldsymbol{n}}$ with the eigenvalue $\pm\frac{\hbar}{2}$. This uses the fact that the eigenvalues of the spin projection along any axis are always $\pm\frac{\hbar}{2}$. We can deduce this by noting that we could have chosen any axis originally as the quantization axis, but you should also try to prove it by expressing a general $\vec{S}\cdot\hat{\boldsymbol{n}}$ as a matrix and computing its eigenvalues.

For example, the eigenkets of S^x are

$$\left|S^x = +\tfrac{\hbar}{2}\right\rangle = \frac{1}{\sqrt{2}}(|+\rangle \pm |-\rangle)\,, \tag{4.6}$$

with eigenvalues $\pm\frac{\hbar}{2}$. We can prove this using the fact that S^x corresponds to the matrix

$$S^x \leftrightarrow \frac{\hbar\sigma^x}{2} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.7}$$

We can then compute

$$\begin{aligned} \text{Prob}\left(S^x = +\frac{\hbar}{2}\right) &= \left|\left\langle S^x = +\tfrac{\hbar}{2}|\psi\right\rangle\right|^2 \\ &= \left|\frac{1}{\sqrt{2}}(\langle +| + \langle -|)(c_+|+\rangle + c_-|-\rangle)\right|^2 \\ &= \left|\frac{1}{\sqrt{2}}(c_+ + c_-)\right|^2 \\ &= \frac{1}{2}|c_+ + c_-|^2\,. \end{aligned} \tag{4.8}$$

4.1.2 The Stern–Gerloch Filter Revisited

Recall the experiment discussed in the first lecture, in which we feed a beam of spin-$\frac{1}{2}$ atoms through a region of spatially-varying magnetic field, resulting in the beam splitting into two beams, one of which has $S^z = +\frac{\hbar}{2}$ and the other of which has $S^z = -\frac{\hbar}{2}$. We discussed several thought experiments using such filters oriented along various axes:

1. First, we discussed feeding the beam through a z-axis filter, and then taking the $S^z = +\frac{\hbar}{2}$ output and feeding it through yet another z-axis filter. The result is that all of the atoms come out of the filter with $S^z = +\frac{\hbar}{2}$. This corresponds to the collapse of the wavefunction; the wavefunction has collapsed after the first measurement, and so the second measurement returns the same result without further altering the wavefunction.

2. Next, we considered feeding the beam through a z-axis filter, and then taking the $S^z = +\frac{\hbar}{2}$ output and feeding it through an x-axis filter. The result is that 50% of the atoms come out with each of $S^x = \pm\frac{\hbar}{2}$. We can see this from the calculation in Eq. (4.8), with $c_+ = 1, c_- = 0$.

3. Finally, we considered feeding the beam through a z-axis filter, feeding the $S^z = +\frac{\hbar}{2}$ output through an x-axis filter, and then feeding the $S^x = +\frac{\hbar}{2}$ output through another z-axis filter. The result is again that 50% of the atoms come out with each of $S^z = \pm\frac{\hbar}{2}$. In this context, we reached the conclusion that we cannot simultaneously measure S^z and S^x. These are incompatible observables.

4.2 Compatible and Incompatible Observables

We have seen that S^z and S^x cannot be simultaneously sharp (just like x and p). When can two observables be simultaneously sharp? Let A, B be two Hermitian operators corresponding to two observables. If we measure A, then the wavefunction collapses to one of the eigenstates $|a\rangle$ of A with eigenvalue a. We next want to measure B and see what happens. Suppose that $|a\rangle$ is also an eigenstate of B with eigenvalue b; in this case, measuring B will return the corresponding eigenvalue b. If we remeasure A, then we will once again find the result a. Any measurements of these two operators will always return sharp values, without changing the state after the first measurement.

Thus, the operators A, B can be simultaneously measured to have sharp values if and only if *all* eigenstates of A are also eigenstates of B. Two such observables are called *compatible*, and two observables not satisfying this relation are called *incompatible*. We can make this statement more concise. If the eigenstates of A are exactly those of B, then A and B are simultaneously diagonalizable, which occurs if and only if $[A, B] = 0$ (we proved this in the last lecture). This is the concise condition for compatibility: two observables are compatible if and only if the corresponding operators commute.

As an example, for the spin-$\frac{1}{2}$ system, the operators $S^2 = S_x^2 + S_y^2 + S_z^2$ and S^z (or the spin along any given axis) are compatible operators, while S^z and S^x are incompatible. Note that

$$[S^z, S^x] = i\hbar S^y \neq 0\,. \tag{4.9}$$

Given a Hilbert space, we can ask what is the maximum number of mutually compatible observables we can find. A *complete set of commuting observables* is a set of observables $\{A, B, C, \dots\}$ such that all pairwise commutators vanish,

$$[A, B] = [A, C] = [B, C] = \cdots = 0\,, \tag{4.10}$$

and such that for any $a, b, c, \ldots$, there is at most one solution to the eigenvalue equation

$$\begin{aligned} A|\alpha\rangle &= a|\alpha\rangle\,, \\ B|\alpha\rangle &= b|\alpha\rangle\,, \\ C|\alpha\rangle &= c|\alpha\rangle\,, \\ &\vdots \end{aligned} \tag{4.11}$$

4.3 The Generalized Uncertainty Relation

Consider two observables A and B. If we measure A, we want to discuss the variance of the probability distribution of possible outcomes. We do this in the standard way: we define the variance of A as

$$\begin{aligned} \Delta A^2 &= \big\langle (A - \langle A\rangle)^2 \big\rangle \\ &= \big\langle A^2 \big\rangle - \langle A\rangle^2 \\ &= \big\langle \psi \big| A^2 \big| \psi \big\rangle - (\langle \psi | A | \psi \rangle)^2\,. \end{aligned} \tag{4.12}$$

If $|\psi\rangle = |a\rangle$ is an eigenket of A with eigenvalue a, then

$$\begin{aligned} \big\langle \psi \big| A^2 \big| \psi \big\rangle &= a^2\,, \\ \langle \psi | A | \psi \rangle &= a\,, \end{aligned} \tag{4.13}$$

so $\Delta A^2 = 0$. In general, however, the variance will not vanish.

We further define the standard deviation $\Delta A = \sqrt{\Delta A^2}$. The uncertainty relation states that

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|\,. \tag{4.14}$$

(Note that this is not the strongest possible form of the uncertainty relation, but it is the most commonly seen form.)

Proof. We will make use of the Cauchy–Schwarz inequality: for all $|a\rangle, |b\rangle \in \mathcal{H}$,

$$|\langle a | b \rangle|^2 \leq \langle a | a \rangle \langle b | b \rangle\,. \tag{4.15}$$

For a general $|\psi\rangle$, define the operators

$$\begin{aligned} \delta A &= A - \langle A\rangle\,, \\ \delta B &= B - |B\rangle\,, \end{aligned} \tag{4.16}$$

and the states

$$\begin{aligned} |a\rangle &= \delta A |\psi\rangle\,, \\ |b\rangle &= \delta B |\psi\rangle\,. \end{aligned} \tag{4.17}$$

Applying the Cauchy–Schwarz inequality, we then have

$$|\langle \psi | \delta A \delta B | \psi \rangle|^2 \leq \big\langle \psi \big| (\delta A)^2 \big| \psi \big\rangle \big\langle \psi \big| (\delta B)^2 \big| \psi \big\rangle\,. \tag{4.18}$$

Note that δA and δB are Hermitian, but there is no guarantee that their product is Hermitian. However, we know that we can write

$$\delta A \delta B = \frac{\delta A \delta B + \delta B \delta A}{2} + \frac{\delta A \delta B - \delta B \delta A}{2}\,. \tag{4.19}$$

The first term on the right-hand side is Hermitian, while the second is anti-Hermitian. We can write this more concisely as

$$\delta A \delta B = \frac{1}{2}\{A, B\} + \frac{1}{2}[A, B]\,. \tag{4.20}$$

Thus,

$$\langle \delta A \delta B \rangle = \frac{1}{2}\langle \{\delta A, \delta B\} \rangle + \frac{1}{2}\langle [\delta A, \delta B] \rangle\,. \tag{4.21}$$

The first term on the right-hand side must be real, as the eigenvalues of any Hermitian operator are real, and the second term on the right-hand side must be imaginary, because any anti-Hermitian operator can be written as i times a Hermitian operator. Thus, the squared modulus of this quantity will simply be the sum of the squared moduli of each of these terms. We then have

$$|\langle \delta A \delta B \rangle|^2 = \frac{1}{4}\big(|\langle \{\delta A, \delta B\} \rangle|^2 + |\langle [\delta A, \delta B] \rangle|^2\big)\,. \tag{4.22}$$

This statement leads to the strongest form of the inequality, but also implies the weaker statement

$$|\langle \delta A \delta B \rangle|^2 \geq \frac{1}{4}|\langle [\delta A, \delta B] \rangle|^2\,. \tag{4.23}$$

Now, note that

$$[\delta A, \delta B] = [A - \langle A \rangle, B - \langle B \rangle] = [A, B]\,, \tag{4.24}$$

so we have

$$|\langle \delta A \delta B \rangle|^2 \geq \frac{1}{4}|\langle [A, B] \rangle|^2\,. \tag{4.25}$$

Returning to Eq. (4.18), we then see that

$$\Delta A^2 \Delta B^2 \geq |\langle \delta A \delta B \rangle|^2 \geq \frac{1}{4}|\langle [A, B] \rangle|^2\,, \tag{4.26}$$

which completes the proof. □

4.4 Position and Momentum

We will now move on to study observables with continuous eigenvalues. In order to do so, we need to work with infinite-dimensional Hilbert spaces. Consider any Hermitian operator ξ with a continuous spectrum

$$\xi \big|\xi'\big\rangle = \xi' \big|\xi'\big\rangle\,, \quad \xi' \in \mathbb{R}\,. \tag{4.27}$$

How should we regard the overlap of two states, $\langle \xi' | \xi'' \rangle$? For a discrete spectrum, we know that the overlap of two distinct (normalized) eigenkets of a Hermitian operator are orthonormal. We want to appropriately generalize this statement. For a discrete spectrum, the orthonormality condition is stated as

$$\big\langle a \big| a' \big\rangle = \delta_{aa'}\,. \tag{4.28}$$

In the continuous case, we generalize this to

$$\big\langle \xi' \big| \xi'' \big\rangle = \delta\big(\xi' - \xi''\big)\,, \tag{4.29}$$

where the Dirac delta "function" $\delta(x)$ is defined to be the object satisfying

$$\int_{-\infty}^{\infty} \mathrm{d}x\; \delta(x) = 1 \tag{4.30}$$

with $\delta(x) = 0$ for all $x \neq 0$. This has the property

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x) f(x) = f(0)\,. \tag{4.31}$$

We can think of this as the limit of a sequence of increasingly peaked functions centered at $x = 0$, each with total area 1.

We will also generalize the complete relation

$$\mathbb{1} = \sum_a |a\rangle\langle a| \tag{4.32}$$

that is familiar from the discrete spectrum case. For continuous spectra, the equivalent statement is

$$\mathbb{1} = \int_{-\infty}^{\infty} \mathrm{d}\xi'\, \big|\xi'\big\rangle\big\langle\xi'\big|\,. \tag{4.33}$$

This is called *resolution of the identity*. For an arbitrary state $|\psi\rangle$, we then have

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}\xi'\, \big|\xi'\big\rangle\big\langle\xi'\big|\psi\big\rangle\,. \tag{4.34}$$

We can express an arbitrary inner product as

$$\big\langle\psi'\big|\psi\big\rangle = \int_{-\infty}^{\infty} \mathrm{d}\xi'\, \big\langle\psi'\big|\xi'\big\rangle\big\langle\xi'\big|\psi\big\rangle\,. \tag{4.35}$$

Lecture 5 (Sep. 20, 2017)

5.1 The Position Operator

In the last class, we talked about operators with a continuous spectrum. A prime example is the position operator. Let's first consider a particle in $d = 1$. We define x as the position operator, with corresponding eigenstates $|x'\rangle$ satisfying

$$x|x'\rangle = x'|x'\rangle\,. \tag{5.1}$$

Here, we are using the Sakurai notation, where operators are symbols with no hats or primes, and the eigenstates are named by their eigenvalues.

We postulate that the $|x'\rangle$ form a complete set of states for the Hilbert space. This means that we can expand any state $|\psi\rangle$ as

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x'\ |x'\rangle\langle x'|\psi\rangle\,. \tag{5.2}$$

In particular, the identity can be resolved as

$$\mathbb{1} = \int_{-\infty}^{\infty} \mathrm{d}x'\ |x'\rangle\langle x'|\,. \tag{5.3}$$

5.1.1 Measurement of Position

Imagine an experiment in which we measure the position of a particle, which could come in the form of a detector that registers a click when a particle enters the detector. Any physical detector will have some finite resolution, below which we cannot resolve the precise location of the particle. Assume we have a detector that detects a particle when it is between $x' - \frac{\Delta}{2}$ and $x' + \frac{\Delta}{2}$.

We will now modify our "collapse" postulate from earlier in the course. After measurement, the state of the particle will be such that a second measurement of x will yield a value between $x' - \frac{\Delta}{2}$ and $x' + \frac{\Delta}{2}$. Mathematically, this means that an initial state

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x'\ |x'\rangle\langle x'|\psi\rangle \tag{5.4}$$

is sent by the measurement to a state

$$|\psi\rangle \to |\psi'\rangle = \int_{x'-\frac{\Delta}{2}}^{x'+\frac{\Delta}{2}} \mathrm{d}x''\ |x''\rangle\langle x''|\psi\rangle\,. \tag{5.5}$$

We assume that the resolution Δ is small by comparison with the length scale of variation of $\langle x''|\psi\rangle$. Then, we can approximate this function to be constant over the range $\left[x' - \frac{\Delta}{2}, x' + \frac{\Delta}{2}\right]$, in which case we find

$$\text{Prob(detection)} = |\langle x'|\psi\rangle|^2 \Delta\,. \tag{5.6}$$

Taking the resolution to be infinitesimal, $\Delta \to \mathrm{d}x'$, this statement becomes

$$\text{Prob(detection)} = |\langle x'|\psi\rangle|^2\, \mathrm{d}x'\,. \tag{5.7}$$

This is the expected statement: the probability density at a point is the squared modulus of the wavefunction at that point. The total probability that the particle exists is

$$\text{Prob(particle is somewhere)} = \int_{-\infty}^{\infty} \mathrm{d}x'\ |\langle x'|\psi\rangle|^2 = 1\,, \tag{5.8}$$

which is true if $\langle\psi|\psi\rangle = 1$. The *position-space wavefunction* is defined as

$$\langle x'|\psi\rangle := \psi(x')\,. \tag{5.9}$$

5.1.2 Hilbert Spaces

We now make a brief aside about the actual definition of a Hilbert space. We generalize our notion of a vector space to the infinite-dimensional case. Consider any sequence of states in the space that has the property that for any $\epsilon > 0$, if we look sufficiently far along in the sequence, any two states in the sequence after that point will be closer together than ϵ under the metric induced by the inner product. The additional property beyond those of a vector space that a Hilbert space has, called completeness, is that all such sequences must converge to another state in the space. In practice, this means that we can approximate any given state arbitrarily well by a series expansion. For more information on Hilbert spaces, see the handwritten notes or the notes from Recitation 1.

5.1.3 Generalizing to Particles in Dimension d

So far we have talked about particles moving in only one dimension. We want to generalize this to particles moving in d dimensions for arbitrary d. In this case, there are d position operators,

$$\vec{x} = (x_1, \ldots, x_d)\,. \tag{5.10}$$

We assert that the $\{x_i\}$ are a set of compatible observables, meaning that we can simultaneously measure each of them. This allows us to know the position of a particle along all axes at the same time. (It is interesting to consider what happens if we relax this assumption, for example, if we have a particle in two dimensions where the x and y operators do not commute. We will discuss such a system later in the course.)

We can then find a complete set of states that are simultaneous eigenstates of each of the position operators, which we label as

$$\vec{x}|\vec{x}'\rangle = \vec{x}'|\vec{x}'\rangle\,, \tag{5.11}$$

and then we proceed as in the $d = 1$ case.

5.2 The Momentum Operator

In the position basis, the momentum operator is given by the familiar expression

$$p_i = -i\hbar\frac{\partial}{\partial x_i}\,. \tag{5.12}$$

On the space of square-integrable functions $\psi(x)$, this operator acts as

$$p\psi(x) = -i\hbar\frac{\mathrm{d}\psi}{\mathrm{d}x}\,. \tag{5.13}$$

From this, we see that

$$[x,p]\psi = -i\hbar\left(x\frac{\mathrm{d}\psi}{\mathrm{d}x} - \frac{\mathrm{d}}{\mathrm{d}x}(x\psi)\right) = i\hbar\psi(x)\,, \tag{5.14}$$

from which we conclude that

$$[x,p] = i\hbar\,. \tag{5.15}$$

Applying the generalized uncertainty principle,

$$\Delta A\Delta B \geq \frac{1}{2}|\langle[A,B]\rangle|\,, \tag{5.16}$$

we find the Heisenberg uncertainty principle,

$$\Delta x\Delta p \geq \frac{\hbar}{2}\,. \tag{5.17}$$

5.2.1 Momentum Basis

The momentum operator is Hermitian, and so we can find a complete set of eigenstates of the momentum operator. The momentum eigenstates $|p'\rangle$ satisfy

$$p|p'\rangle = p'|p'\rangle\,, \tag{5.18}$$

and they form an orthonormal basis with

$$\langle p''|p'\rangle = \delta(p'' - p')\,. \tag{5.19}$$

The identity can then be resolved as

$$\mathbb{1} = \int_{-\infty}^{\infty} \mathrm{d}p'\ |p'\rangle\langle p'|\,, \tag{5.20}$$

which allows us to decompose an arbitrary state $|\psi\rangle$ as

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p'\ |p'\rangle\langle p'|\psi\rangle\,. \tag{5.21}$$

Analogously to the discussion of measurement of position, measurements of momentum will have

$$\mathrm{Prob}\big(\text{momentum is between } p' \text{ and } p' + \mathrm{d}p'\big) = \big|\langle p'|\psi\rangle\big|^2\,\mathrm{d}p'\,. \tag{5.22}$$

We similarly define the *momentum-space wavefunction* as

$$\langle p'|\psi\rangle := \psi(p')\,. \tag{5.23}$$

Although we will rarely use this notation, it is worth noting that Sakurai denotes the momentum-space wavefunction as

$$\langle p'|\psi\rangle := \phi(p')\,. \tag{5.24}$$

How are $|p'\rangle$ and $|x'\rangle$ related? We need to find $\langle x'|p'\rangle$. We can determine this by inserting a momentum operator:

$$-i\hbar\frac{\partial}{\partial x}\langle x'|p'\rangle = \langle x'|p|p'\rangle = p'\langle x'|p'\rangle\,. \tag{5.25}$$

The expression on the left-hand side is found by using the expression for the momentum operator in the position basis,since $\langle x'|p'\rangle$ is the position-space wavefunction for the state $|p'\rangle$, while the expression on the right-hand side is found by having the momentum operator act on the ket $|p'\rangle$. This solution to this differential equation is

$$\langle x'|p'\rangle = Ne^{ip'x'/\hbar} \tag{5.26}$$

for some normalization factor N.

To fix the normalization, we use

$$\langle x'|x''\rangle = \delta(x' - x'')\,, \tag{5.27}$$

which implies

$$\int_{-\infty}^{\infty} \mathrm{d}p'\ \langle x'|p'\rangle\langle p'|x''\rangle = \delta(x' - x'')\,. \tag{5.28}$$

The left-hand side of this equation is

$$\int_{-\infty}^{\infty} \mathrm{d}p' \, \langle x'|p'\rangle\langle p'|x''\rangle = |N|^2 \int_{-\infty}^{\infty} \mathrm{d}p' \, e^{ip'(x'-x'')/\hbar} = 2\pi\hbar|N|^2\delta\big(x'-x''\big)\,, \tag{5.29}$$

which must equal the right-hand side,

$$2\pi\hbar|N|^2\delta\big(x'-x''\big) = \delta\big(x'-x''\big)\,. \tag{5.30}$$

This only fixes the modulus of N, but by convention we choose N to be real and positive, so that we have

$$N = \frac{1}{\sqrt{2\pi\hbar}}\,. \tag{5.31}$$

This finally gives us

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ip'x'/\hbar}\,. \tag{5.32}$$

This allows us to relate the wavefunctions $\langle x'|\psi\rangle$ and $\langle p'|\psi\rangle$ in the position and momentum bases. By inserting the identity, we can write

$$\langle x'|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p' \, \langle x'|p'\rangle\langle p'|\psi\rangle = \int_{-\infty}^{\infty} \frac{\mathrm{d}p'}{\sqrt{2\pi\hbar}} \, e^{ip'x'/\hbar}\langle p'|\psi\rangle\,. \tag{5.33}$$

Thus, we see that the momentum-space wavefunction is related to the position space wavefunction by the Fourier transform. We similarly have

$$\langle p'|\psi\rangle = \int \frac{\mathrm{d}x'}{\sqrt{2\pi\hbar}} \, e^{-ip'x'/\hbar}\langle x'|\psi\rangle\,. \tag{5.34}$$

5.3 Normalization of Position and Momentum Eigenstates

What is the wavefunction corresponding to the position eigenstate $|x'\rangle$? We have

$$\langle x''|x'\rangle = \delta\big(x''-x'\big)\,. \tag{5.35}$$

This is not a square-integrable function, and so it is not actually an element of the Hilbert space, but we can nonetheless build all square-integrable functions out of the position states. Similarly, the momentum eigenstates are not elements of the Hilbert space, but can be used to decompose any square-integrable function.

How can we make this mathematically precise? We do this by changing the problem, and stating that our actual problem of interest is well-approximated by the new problem. The states $|x'\rangle$ and $|p'\rangle$ are not normalizable, and so we will change the problem to introduce normalizable states. We do this by assuming that our particle is moving in some finite-size space, and then take the finite size to infinity at the end of the problem.

For example, we can define momentum states by putting our particle in a large box of size L with periodic boundary conditions,

$$e^{ip'(x'+L)/\hbar} = e^{ip'x'/\hbar}\,. \tag{5.36}$$

This implies that

$$p'L = 2\pi n\,, \quad n \in \mathbb{Z}\,. \tag{5.37}$$

We now have discrete momenta, and a countable basis of momentum eigenstates. For finite L, we require that

$$\int_0^L \mathrm{d}x' \, \left|\langle x'|p'\rangle\right|^2 = 1\,, \tag{5.38}$$

which normalizes the momentum eigenstates:

$$\langle x'|p'\rangle = \frac{1}{\sqrt{L}} e^{ip'x'/\hbar}\,. \tag{5.39}$$

5.4 The Uncertainty Principle

The uncertainty relation tells us that if the position-space wavefunction $\psi(x')$ has a shape with some width ℓ, then the momentum-space wavefunction $\phi(p')$ will have a shape with some width of order $\hbar/\ell$, as seen in Figure 1.

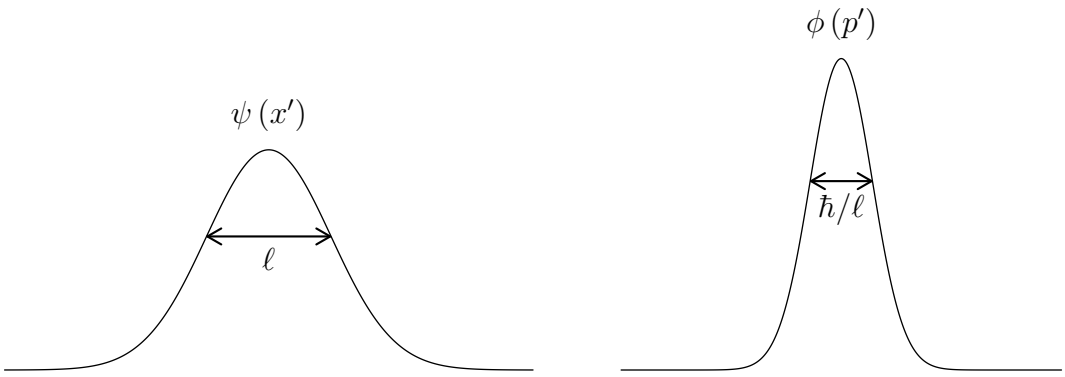


Figure 1: Example position-space and momentum-space wavefunctions that are related by the Fourier transform. The uncertainty relation tells us that their widths will be roughly reciprocal to one another.

5.4.1 Minimum Uncertainty States

For what wavefunction is the uncertainty relation $\Delta x \Delta p \geq \frac{\hbar}{2}$ saturated? The answer is a simple Gaussian,

$$\psi(x') = A e^{-\frac{x'^2}{2w}}\,. \tag{5.40}$$

You will prove this in the homework.

5.5 Momentum and Translation

We will now use the momentum operator p to define a unitary operator

$$T(a) = e^{-iap/\hbar}\,, \quad a \in \mathbb{R}\,. \tag{5.41}$$

This operator is unitary because p is Hermitian:

$$(T(a))^\dagger = e^{iap/\hbar} = (T(a))^{-1}\,. \tag{5.42}$$

The operator exponential used here can be defined using the series expansion,

$$T(a) = e^{-iap/\hbar} = \sum_{n=0}^{\infty} \frac{(-ia/\hbar)^n p^n}{n!} \,. \tag{5.43}$$

Note that this definition can only make sense when acting on a Hilbert space.

We can see that this operator satisfies the following properties:

1. $T^{-1}(a) = T(-a)$
2. $T(a')T(a'') = T(a' + a'')$
3. $T(a)xT^{-1}(a) = x - a$

Let's prove the final property.

Proof. Define

$$F(a) = T(a)xT^{-1}(a) = e^{-ipa/\hbar} x e^{ipa/\hbar} \,. \tag{5.44}$$

The trick to the proof will be to compute the derivative of this quantity with respect to a:

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}a} &= \frac{i}{\hbar}\Big(e^{-ipa/\hbar}(-px)e^{ipa/\hbar} + e^{-ipa/\hbar}xpe^{ipa/\hbar}\Big) \\ &= \frac{i}{\hbar} e^{-ipa/\hbar}[x,p,e]^{ipa/\hbar} \\ &= \frac{i}{\hbar} e^{-ipa/\hbar} i\hbar e^{ipa/\hbar} \\ &= -1 \,. \end{aligned} \tag{5.45}$$

We can then integrate this equation to find

$$F(a) = F(0) - a \,. \tag{5.46}$$

Because $F(0) = x$, this completes the proof. □

For this reason, we will call $T(a)$ the *translation operator*. What happens to a state $|\psi\rangle$ under the action of this operator? We know that

$$T(a)xT^{-1}(a)\big|x'\big\rangle = (x-a)\big|x'\big\rangle = \big(x'-a\big)\big|x'\big\rangle \,. \tag{5.47}$$

By applying $T^{-1}(a)$ to each side of this equation, we find

$$xT^{-1}(a)|x'\rangle = \big(x'-a\big)T^{-1}(a)\big|x'\big\rangle \,, \tag{5.48}$$

from which we conclude that

$$T^{-1}(a)\big|x'\big\rangle \propto \big|x'-a\big\rangle \,. \tag{5.49}$$

This operator does not change the modulus of a state because it is unitary, so we actually have

$$T^{-1}(a)\big|x'\big\rangle = \big|x'-a\big\rangle \,. \tag{5.50}$$

The effect on the wavefunction is thus

$$\big\langle x'\big|T(a)\big|\psi\big\rangle = (T^{-1}(a)\big|x'\big\rangle, |\psi\rangle) = \big\langle x'-a\big|\psi\big\rangle = \psi\big(x'-a\big) \,. \tag{5.51}$$

We see that $T(a)|\psi\rangle$ has the shifted wavefunction $\psi(x'-a)$ if $|\psi\rangle$ has the wavefunction $\psi(x')$. This wavefunction is shifted to the right.

We have now seen that this unitary translation operator, which is a natural operator to define acting on the Hilbert space, is determined entirely be the Hermitian operator p. We could have turned this story around, by first defining the unitary translation operator and then using it to define a Hermitian operator p. In the classical limit, the operator p found in this way has the right properties to be called a momentum. Note that any unitary operator U can be written as e^{iA} for some Hermitian operator A. This is the approach we will take when we discuss angular momentum later in the course: it will be defined to be the operator that, when exponentiated, rotates the quantum mechanical system.

Lecture 6 (Sep. 25, 2017)

6.1 Solving Problems in Convenient Bases

Last time we talked about the momentum and position bases. From a practical point of view, it often pays off to not solve a particular problem in the position basis, and instead choose a more convenient basis, such as the momentum basis. Suppose we have a Hamiltonian

$$H = -\frac{1}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + ax\,. \tag{6.1}$$

We could solve this in the position basis; it is a second-order differential equation, with solutions called Airy functions. However, this is easy to solve in the momentum basis. In the momentum basis, the Hamiltonian becomes

$$H = \frac{p^2}{2m} + ia\hbar\frac{\mathrm{d}}{\mathrm{d}p}\,, \tag{6.2}$$

which is only a first-order differential equation.

For other problems, the most convenient basis may not be the position or the momentum basis, but rather some mixed basis. It is worth becoming comfortable with change of basis in order to simplify problems.

6.2 Quantum Dynamics

Recall the fourth postulate: time evolution is a map

$$|\psi(t)\rangle \mapsto \big|\psi\big(t'\big)\big\rangle = U\big(t',t\big)|\psi(t)\rangle\,, \tag{6.3}$$

where $U(t',t)$ is the time-evolution operator, which is unitary. In the case where t' and t differ by an infinitesimal amount, we can rewrite time evolution as a differential equation,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle\,, \tag{6.4}$$

where $H(t)$ is the Hamiltonian, which is a Hermitian operator.

In general, we can always write

$$\big|\psi\big(t'\big)\big\rangle = U\big(t',t\big)|\psi(t)\rangle \tag{6.5}$$

for *some* operator $U(t',t)$. What statements can we make about the operator $U(t',t)$? We can say the following:

1. The probability must be conserved. We start with a state for which the total probability of all outcomes of a measurement is 1, and so we expect that after evolving in time, we should arrive at a state that also has total probability 1. This implies that U must be unitary, which we will now prove.

   Proof. Let us expand the state $|\psi(t)\rangle$ as

   $$|\psi(t)\rangle = \sum_n c_n(t)|a_n\rangle\,, \tag{6.6}$$

with $\{|a_n\rangle\}$ a basis of eigenstates of any observable A. The probability of a measurement of A yielding the result a_n at time t is given by

$$\text{Prob}(A = a_n \text{ at time } t) := p(a_n; t) = |c_n(t)|^2 \,, \tag{6.7}$$

where we have introduced the shorthand notation $p(a_n; t)$ for this probability. The total probability must be 1,

$$\sum_n p(a_n; t) = 1 \,, \tag{6.8}$$

which is simply the statement that the state $|\psi(t)\rangle$ must be normalized,

$$\sum_n |c_n(t)|^2 = 1 \,. \tag{6.9}$$

At a later time t', we have

$$p\big(a_n; t'\big) = \big|c_n\big(t'\big)\big|^2 \,, \tag{6.10}$$

and once again, requiring that the total probability be 1 requires that $|\psi(t')\rangle$ be normalized,

$$\sum_n \big|c_n\big(t'\big)\big|^2 = 1 \,. \tag{6.11}$$

Thus, we must have

$$\big\langle \psi\big(t'\big) \big| \psi\big(t'\big) \big\rangle = \Big\langle \psi(t) \Big| U^\dagger\big(t', t\big) U\big(t', t\big) \Big| \psi(t) \Big\rangle = 1 \tag{6.12}$$

for arbitrary normalized $|\psi(t)\rangle$. This is satisfied if and only if $U^\dagger(t', t)U(t', t) = \mathbb{1}$, i.e., if U is unitary. □

2. We demand that the time-evolution operator satisfy composition, i.e.,

$$U(t_f, t_i) = U(t_f, t)U(t, t_i) \tag{6.13}$$

for all $t_i < t < t_f$.

3. We require that

$$U(t, t) = \mathbb{1} \,. \tag{6.14}$$

We will now build up an arbitrary time-evolution operator as a sequence of infinitesimal time-evolution operators. Consider a time evolution $U(t + \mathrm{d}t, t)$; we can expand this as a power series in the infinitesimal $\mathrm{d}t$, yielding

$$U(t + \mathrm{d}t, t) = \mathbb{1} - \frac{i}{\hbar} H \, \mathrm{d}t + O\big((\mathrm{d}t)^2\big) \tag{6.15}$$

for some operator H (which we do not yet know anything about).The factors of i and $\hbar$ are put in for historical reasons, but note that we can always expand $U(t + \mathrm{d}t, t)$ in this way. We will neglect higher-order terms because $\mathrm{d}t$ is infinitesimal. We then have

$$U^\dagger(t + \mathrm{d}t, t) = \mathbb{1} + \frac{i}{\hbar} H^\dagger \, \mathrm{d}t + O\big((\mathrm{d}t)^2\big) \,. \tag{6.16}$$

Requiring unitarity, $U^\dagger U = \mathbb{1}$, then gives us

$$U^\dagger U = \left(\mathbb{1} - \frac{i}{\hbar} H\,\mathrm{d}t + \cdots\right)\left(\mathbb{1} + \frac{i}{\hbar} H^\dagger\,\mathrm{d}t + \cdots\right) = \mathbb{1} + \frac{i}{\hbar}\Big(H^\dagger - H\Big)\,\mathrm{d}t + O\big((\mathrm{d}t)^2\big) = \mathbb{1}\,. \quad (6.17)$$

We conclude that we must have $H^\dagger = H$, i.e., H must be Hermitian. The statement that U is a unitary operator is equivalent to the statement that H is Hermitian. We will call H the *Hamiltonian*. Shortly, we will make the connection between this operator and the Hamiltonian we know from classical physics.

Applying the infinitesimal time-evolution operator to a state, we find

$$|\psi(t + \mathrm{d}t, t)\rangle = \left(\mathbb{1} - \frac{i}{\hbar} H\,\mathrm{d}t + O\big((\mathrm{d}t)^2\big)\right)|\psi(t)\rangle\,. \quad (6.18)$$

We can write this statement as

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle\,, \quad (6.19)$$

which we recognize as the Schrödinger equation.

Now let's connect this to the discussion of unitary spatial translations. We can think of time evolution as translation in time. At the end of the last lecture, we alluded to the fact that an infinitesimal spatial translation is generated by the Hermitian momentum operator. This is exactly analogous to the statement that infinitesimal time evolutions are generated by the Hamiltonian, which is Hermitian. (Note that momentum is the conserved Noether charge associated to spatial translations, and energy is the conserved Noether charge associated to time translations.)

6.2.1 Brief Aside on Classical Mechanics

In classical mechanics, time evolution is dictated by Newton's laws. Over the centuries, Newton's laws have been massaged into various forms. Classically, the state of a particle is specified by its position $q(t)$ and momentum $p(t)$. Given such a state, how do $q(t), p(t)$ change as a function of time? This is described by the equations of motion for the particle, which are known as Hamilton's equations:

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}\,, \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}\,, \quad (6.20)$$

where $H = H(q, p)$ is the classical Hamiltonian.

As an example, consider

$$H(q, p) = \frac{p^2}{2m} + V(q)\,. \quad (6.21)$$

Hamilton's equations then give

$$\dot{q} = \frac{p}{m}\,, \quad \dot{p} = -\nabla_q V\,, \quad (6.22)$$

which are simply Newton's equations of motion.

The space of (q, p) is known as *phase space*. The classical time evolution given by Hamilton's equation is such that the area of a region of phase space is preserved by the time evolution (Liouville's theorem). Suppose we have a collection of particles, each with different values of q, p. This collection fills out a region in phase space. If we evolve each particle in time, the shape of the region in phase space will change, but its area will be conserved.

We can state Liouville's theorem formally in terms of the Poisson bracket, defined as

$$\{f, g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial g}{\partial q}\frac{\partial f}{\partial p}\,. \quad (6.23)$$

Any observable A (that only depends on time through q and p) evolves in time according to

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \{A, H\}\,. \tag{6.24}$$

In particular, taking $A = q$ yields

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \{q, H\} = \frac{\partial H}{\partial p} \tag{6.25}$$

and taking $A = p$ yields

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \{p, H\} = -\frac{\partial H}{\partial q}\,. \tag{6.26}$$

Classically, the Hamiltonian is (usually) the total energy of the system. In quantum mechanics, we refer to H as the Hamiltonian because it is the operator that returns the energy of the system. In quantum mechanics, the Hamiltonian (energy) operator and the time operator are more deeply related because of the time-energy uncertainty relation.

6.2.2 The Heisenberg Picture

Let's start with the Schrödinger equation

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle\,. \tag{6.27}$$

For a closed system, H is independent of time, and we can solve this equation exactly, in some formal sense, as

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle\,. \tag{6.28}$$

Thus, the time-evolution operator is given by

$$U(t,0) = e^{-iHt/\hbar}\,. \tag{6.29}$$

The physically significant quantities are matrix elements of operators,

$$\langle\psi'(t)|A|\psi(t)\rangle = \left\langle\psi'(0)\middle|U^\dagger(t,0)AU(t,0)\middle|\psi(0)\right\rangle. \tag{6.30}$$

Here, we have assumed that A does not have explicit time dependence. Thus, the time dependence of the matrix element comes only from the fact that the states change in time. However, we can look at the expression on the right-hand side and define a time-dependent operator

$$A_\mathrm{H}(t) := U^\dagger(t,0)A_\mathrm{H}(0)U(t,0)\,, \quad A_\mathrm{H}(0) = A\,, \tag{6.31}$$

that acts on states $|\psi_H\rangle := |\psi(0)\rangle$ that do not evolve in time. Viewing time evolution in this way is referred to as working in the *Heisenberg picture*. Up until now, we have been working in the *Schrödinger picture*, in which the time dependence of matrix elements comes from the time dependence of states $|\psi(t)\rangle$, and operators without explicit time dependence do not evolve in time. By design, we have

$$\langle\psi'_\mathrm{H}|A_\mathrm{H}(t)|\psi_\mathrm{H}\rangle = \left\langle\psi'(0)\middle|U^\dagger(t,0)AU(t,0)\middle|\psi(0)\right\rangle. \tag{6.32}$$

We can write an infinitesimal version of Eq. (6.31) as

$$\begin{aligned}\frac{\mathrm{d}A_\mathrm{H}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{iHt/\hbar}A_\mathrm{H}(0)e^{-iHt/\hbar}\right)\\ &= \frac{i}{\hbar}e^{iHt/\hbar}(HA_\mathrm{H}(0) - A_\mathrm{H}(0)H)e^{-iHt/\hbar}\\ &= \frac{1}{i\hbar}[A_\mathrm{H}(t), H]\,.\end{aligned} \tag{6.33}$$

Comparing this with the classical time-evolution expression

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \{A, H\}\,, \tag{6.34}$$

we see that the commutator in quantum mechanics plays a similar role as the Poisson bracket in classical mechanics.

6.3 Energy Eigenstates

We will now assume that the Hamiltonian H corresponds to the energy of the system. Because H is Hermitian, we can diagonalize it with a unitary transformation. Let $|j\rangle$ be its eigenstates,

$$H|j\rangle = E_j|j\rangle\,, \tag{6.35}$$

which form an orthonormal basis. We call the $|j\rangle$ *energy eigenstates*, and the E_j the corresponding *energy eigenvalues*. An arbitrary state $|\psi\rangle$ can be decomposed as

$$|\psi\rangle = \sum_j c_j|j\rangle \tag{6.36}$$

with

$$c_j = \langle j|\psi\rangle\,. \tag{6.37}$$

Similarly, an arbitrary operator A can be decomposed as

$$A = \sum_{j,j'} A_{jj'}\big|j\big\rangle\big\langle j'\big|\,. \tag{6.38}$$

In the Schrödinger picture, $A_{jj'}$ is time-independent, and the states evolve as

$$\begin{aligned} |\psi(0)\rangle \mapsto |\psi(t)\rangle &= e^{-iHt/\hbar}|\psi(0)\rangle \\ &= e^{-iHt/\hbar}\sum_j c_j|j\rangle \\ &= \sum_j c_j e^{-iE_jt/\hbar}|j\rangle\,. \end{aligned} \tag{6.39}$$

Thus, we can write

$$|\psi(t)\rangle = \sum_j c_j(t)|j\rangle\,, \tag{6.40}$$

where the coefficients $c_j(t)$ evolve in time as

$$c_j(t) = c_j e^{-iE_jt/\hbar}\,. \tag{6.41}$$

The matrix element of A is then

$$\begin{aligned} \langle\psi(t)|A|\psi(t)\rangle &= \sum_{j,j'} c_j^*(t)c_{j'}(t)\langle j|A|j'\rangle \\ &= \sum_{j,j'} c_j^*(0)c_{j'}(0)e^{i(E_j-E_{j'})t/\hbar}A_{jj'}\,. \end{aligned} \tag{6.42}$$

In the Heisenberg picture, the states are time-independent, so the c_j are fixed, while A evolves as

$$\begin{aligned} A \mapsto A(t) &= e^{iHt/\hbar} A e^{-iHt/\hbar} \\ &= \sum_{j,j'} A_{jj'} \left(e^{iHt/\hbar} |j\rangle\langle j'| e^{-iHt/\hbar} \right) \\ &= \sum_{j,j'} \left(A_{jj'} e^{i(E_j - E_{j'})t/\hbar} \right) |j\rangle\langle j'| \,. \end{aligned} \tag{6.43}$$

The matrix element is then

$$\langle\psi|A(t)|\psi\rangle = \sum_{j,j'} c_j^* c_{j'} e^{i(E_j - E_{j'})t/\hbar} A_{jj'} \,, \tag{6.44}$$

which agrees with the result in the Schrödinger picture.

6.4 Example: Spin Precession In a Magnetic Field

Consider a spin-$\frac{1}{2}$ moment in a magnetic field, with Hamiltonian

$$H = -\frac{ge}{2m} \boldsymbol{S} \cdot \boldsymbol{B} \,. \tag{6.45}$$

Here, e is the electric charge, m is the mass of the particle, and g is the g-factor, which has $g \approx 2$ for the electron. Assuming that $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, we can write the Hamiltonian as

$$H = -\frac{ge}{2m} S^z B \,. \tag{6.46}$$

With this assumption, the S^z eigenstates are also energy eigenstates,

$$H|S^z_\pm\rangle = \mp \frac{ge\hbar B}{4m} |S^z_\pm\rangle \tag{6.47}$$

with energy eigenvalues

$$E_\pm = \mp \frac{ge\hbar B}{4m} \,. \tag{6.48}$$

The energy separation is

$$\hbar\omega = |E_+ - E_-| = \left| \frac{ge\hbar B}{2m} \right| . \tag{6.49}$$

Lecture 7 (Sep. 27, 2017)

7.1 Spin Precession In a Magnetic Field

Last time, we began discussing the classic example of precession of a spin-$\frac{1}{2}$ particle in a magnetic field. The Hamiltonian of this system is

$$H = -\frac{ge}{2m}\boldsymbol{S}\cdot\boldsymbol{B}\,. \tag{7.1}$$

With $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, this becomes

$$H = -\frac{ge}{2m}S^z B\,. \tag{7.2}$$

The energy levels are

$$E_\pm = \mp\frac{ge\hbar B}{4m}\,, \tag{7.3}$$

giving a level splitting of

$$\hbar\omega = \left|\frac{ge\hbar B}{2m}\right|. \tag{7.4}$$

We define

$$\omega := \frac{g|eB|}{2m}\,, \tag{7.5}$$

so that the Hamiltonian can be written simply as

$$H = \omega S^z\,. \tag{7.6}$$

7.1.1 Schrödinger Picture

The time-evolution operator can be expressed in the energy eigenbasis (which coincides with the S^z eigenbasis) as

$$U(t,0) = e^{-iHt/\hbar} = e^{-i\omega S^z t/\hbar}\,. \tag{7.7}$$

Suppose that we have an arbitrary initial state

$$|\psi\rangle = c_+|+\rangle + c_-|-\rangle\,. \tag{7.8}$$

Using the time-evolution operator, we find that

$$\begin{aligned}|\psi(t)\rangle &= U(t,0)|\psi\rangle\\ &= e^{-i\omega S^z t/\hbar}|\psi\rangle\\ &= e^{-i\omega t/2}c_+|+\rangle + e^{i\omega t/2}c_-|-\rangle\,.\end{aligned} \tag{7.9}$$

We now know the state of the system at all times. For example, if we initially have $|\psi\rangle = |+\rangle$, then

$$|\psi(t)\rangle = e^{-i\omega t/2}|+\rangle\,, \tag{7.10}$$

which has

$$\mathrm{Prob}\big(S^z = \tfrac{\hbar}{2}\big) = 1 \tag{7.11}$$

for all times. This is why energy eigenstates are often called *stationary states*. Because time evolution is generated by the Hamiltonian, energy eigenstates do not change under time evolution (up to an unphysical overall phase).

Consider instead the case where the initial state is the spin-up eigenstate of S^x, i.e.,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)\,. \tag{7.12}$$

Then we have

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\Big(e^{-i\omega t/2}|+\rangle + e^{i\omega t/2}|-\rangle\Big)\,. \tag{7.13}$$

Thus, after a time t we have

$$\begin{aligned}\text{Prob}\big(S^x = \tfrac{\hbar}{2} \text{ at time } t\big) &= |\langle S^x_+|\psi(t)\rangle|^2 \\ &= \frac{1}{2}|(\langle +| + \langle -|)|\psi(t)\rangle|^2 \\ &= \cos^2\left(\frac{\omega t}{2}\right).\end{aligned} \tag{7.14}$$

Similarly,

$$\text{Prob}\big(S^x = -\tfrac{\hbar}{2} \text{ at time } t\big) = \sin^2\left(\frac{\omega t}{2}\right). \tag{7.15}$$

We can check that this is true, but we know it must be true by conservation of probability. We then have

$$\langle S^x\rangle = \frac{\hbar}{2}\left[\cos^2\left(\frac{\omega t}{2}\right) - \sin^2\left(\frac{\omega t}{2}\right)\right] = \frac{\hbar}{2}\cos(\omega t)\,. \tag{7.16}$$

We see that this oscillates as a function of time, with angular frequency ω, as we expect from our classical intuition. What can we learn from this calculation? If we have a general initial state represented by $\hat{\boldsymbol{n}}$ on the Bloch sphere, then $\hat{\boldsymbol{n}}$ will precess around $\boldsymbol{B}$ with angular frequency ω. You will show this on the homework.

7.1.2 Heisenberg Picture

We can carry out this same calculation in the Heisenberg picture. In this picture, the states do not change as a function of time, but rather the operators do. The spin operator in the Heisenberg picture evolves according to the equation

$$\boldsymbol{S}(t) = e^{i\omega S^z t/\hbar}\boldsymbol{S}e^{-i\omega S^z t/\hbar}\,. \tag{7.17}$$

Taking the time derivative of this expression, we find

$$\frac{\mathrm{d}\boldsymbol{S}}{\mathrm{d}t} = \frac{i\omega}{\hbar}e^{i\omega S^z t/\hbar}[S^z, \boldsymbol{S}]e^{-i\omega S^z t/\hbar}\,. \tag{7.18}$$

The z-component of this equation of motion is simple, because $[S^z, S^z] = 0$, so we have

$$\frac{\mathrm{d}S^z}{\mathrm{d}t} = 0\,. \tag{7.19}$$

The x-component is found using the fact that $[S^z, S^x] = i\hbar S^y$, which gives

$$\frac{\mathrm{d}S^x}{\mathrm{d}t} = \frac{i\omega}{\hbar}e^{i\omega S^z t/\hbar}i\hbar S^y e^{-i\omega S^z t/\hbar} = -\omega S^y(t)\,. \tag{7.20}$$

Similarly, using $[S^z, S^y] = -i\hbar S^x$, we find

$$\frac{\mathrm{d}S^y}{\mathrm{d}t} = \omega S^x(t)\,. \tag{7.21}$$

We can write these three expression compactly in vector notation as

$$\frac{\mathrm{d}\boldsymbol{S}}{\mathrm{d}t} = -\omega \boldsymbol{S} \times \hat{\boldsymbol{z}} \,. \tag{7.22}$$

This is the same equation as for classical spins, but the interpretation is entirely different because $\boldsymbol{S}$ is now an operator. If we take the expectation value of each side of this expression, we will get the same answers that we found in the Schrödinger picture.

7.2 Particle in a Potential

We have said, on general grounds, that time evolution is given by a unitary operator. Furthermore, we have said that the infinitesimal form of this operator corresponds to a Hermitian operator, which we have called the Hamiltonian. However, we must ask: how do we specify a quantum mechanical system? There are two main ingredients we need: we have to first specify the Hilbert space that the states live in, and then we have to specify the Hamiltonian. Either of these is insufficient without the other.

For a particle in a potential, the Hilbert space is the space of square-integrable functions (modulo magnitude and phase). In order to specify the Hamiltonian, we will simply take the classical Hamiltonian and replace the variables x and p by the corresponding operators x and p. We must be careful, because although the regular variables x and p commute, the corresponding operators do not.

Suppose we have the Hamiltonian

$$H = \frac{p^2}{2m} + V(x) \,. \tag{7.23}$$

How do we proceed? In the Schrödinger picture, we first find the energy eigenkets $|j\rangle$ and eigenvalues E_j, which satisfy

$$H|j\rangle = E_j|j\rangle \,. \tag{7.24}$$

Once we have found these eigenkets, we can expand an arbitrary state as

$$|\psi\rangle = \sum_j c_j |j\rangle \,. \tag{7.25}$$

Time evolution is carried out as

$$|\psi(t)\rangle = \sum_j c_j e^{-iE_j t/\hbar} |j\rangle \,. \tag{7.26}$$

In principle, this is a complete procedure to determine the dynamics of any quantum system. However, in practice this can often be very difficult.

In the Heisenberg picture, we simply care about the operators,

$$\begin{aligned} x_{\mathrm{H}}(t) &= e^{iHt/\hbar} x e^{-iHt/\hbar} \,, \\ p_{\mathrm{H}}(t) &= e^{iHt/\hbar} p e^{-iHt/\hbar} \,. \end{aligned} \tag{7.27}$$

Using the Heisenberg picture equation of motion, we see that the operator x_H evolves according to

$$\frac{\mathrm{d}x_{\mathrm{H}}}{\mathrm{d}t} = \frac{1}{i\hbar}[x_{\mathrm{H}}, H] = \frac{1}{i\hbar}\left[x_{\mathrm{H}}, \frac{p_{\mathrm{H}}^2}{2m} + V(x_{\mathrm{H}})\right] = \frac{1}{i\hbar}\left[x_{\mathrm{H}}, \frac{p_{\mathrm{H}}^2}{2m}\right] . \tag{7.28}$$

From this point forward, we will drop the subscript H on the Heisenberg picture operators. Using a commutator identity, we then have

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{1}{i\hbar}\left[x, \frac{p^2}{2m}\right] \\ &= \frac{1}{i\hbar}\left(\left[x, \frac{p}{2m}\right]p + p\left[x, \frac{p}{2m}\right]\right) \\ &= \frac{p}{m}\,.\end{aligned} \tag{7.29}$$

This is the expected result, but it is now an operator equation. Similarly, we find

$$\begin{aligned}\frac{\mathrm{d}p}{\mathrm{d}t} &= \frac{1}{i\hbar}[p, V(x)] \\ &= \frac{1}{i\hbar}\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}(V(x)\cdot) + V(x)i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\cdot\right) \\ &= -\frac{\mathrm{d}V}{\mathrm{d}x}\,.\end{aligned} \tag{7.30}$$

Here, we have made clear the action of the derivative operator by using $\cdot$ to denote an arbitrary position-space wavefunction on which these operators could act. Once again, this resembles the classical expression, but is now an operator equation.

If we take the expectation values of Eqs. (7.29) and (7.30), we find

$$\begin{aligned}\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} &= \frac{\langle p\rangle}{m}\,, \\ \frac{\mathrm{d}\langle p\rangle}{\mathrm{d}t} &= -\left\langle\frac{\mathrm{d}V}{\mathrm{d}x}\right\rangle\,.\end{aligned} \tag{7.31}$$

This result is called *Ehrenfest's theorem*. It is important, when using the second equation, to remember to differentiate the potential first before taking the expectation value, because reversing the order of these operators will often change the result.

7.2.1 Example: Charged Particle in a Uniform Electric Field

Consider a charged particle in a uniform electric field, which has the Hamiltonian

$$H = \frac{p^2}{2m} - qE(t)x\,. \tag{7.32}$$

In the Schrödinger picture, this is a messy problem to solve. In the Heisenberg picture, however, the problem is not difficult at all. Using Eq. (7.30), we have

$$\frac{\mathrm{d}p}{\mathrm{d}t} = qE(t)\,. \tag{7.33}$$

This has solution

$$p(t) = p(0) + \int_0^t \mathrm{d}t'\; qE\big(t'\big)\,. \tag{7.34}$$

Similarly, using Eq. (7.29), we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{p(t)}{m}\,, \tag{7.35}$$

which yields

$$x(t) - x(0) = \frac{tp(0)}{m} + \frac{q}{m}\int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' \, E\big(t''\big) \,. \tag{7.36}$$

These are the same results we find classically. If $E(t) = E$ is independent of time, then we have

$$p(t) = p(0) + qEt \,, \quad x(t) = x(0) + \frac{p(0)t}{m} + \frac{q}{2m}Et^2 \,. \tag{7.37}$$

7.2.2 Example: Simple Harmonic Oscillator

Recall the simple harmonic oscillator Hamiltonian,

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \,. \tag{7.38}$$

We are likely all familiar with the approach in the Schrödinger picture. In the Heisenberg picture, we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{p}{m} \,, \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -m\omega^2 x \,. \tag{7.39}$$

Solving these equations gives

$$\begin{aligned} x(t) &= x(0)\cos(\omega t) + \frac{p(0)}{m\omega}\sin(\omega t) \,, \\ p(t) &= -m\omega x(0)\sin(\omega t) + p(0)\cos(\omega t) \,. \end{aligned} \tag{7.40}$$

Finding these equations in the Schrödinger picture is messy, though it can be done; in the Heisenberg picture, the result was immediate.

Keep in mind that these are operator equations. If, for example, we square the operators $x(t)$ or $p(t)$, we must be careful with the order of $x(0)$ and $p(0)$ in the cross terms, as these operators do not commute.

Lecture 8 (Oct. 2, 2017)

8.1 General Time Dependent Hamiltonians

The Schrödinger equation dictates that quantum states evolve in time according to

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle\,. \tag{8.1}$$

In the last class, we saw that if the Hamiltonian is independent of time, $H(t) = H$, then we can solve this differential equation as

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle\,. \tag{8.2}$$

We are now interested in what happens if H has explicit time dependence.

We can always expand an infinitesimal time-evolution operator as

$$U(t+\mathrm{d}t,t) = 1 - \frac{i}{\hbar}H(t)\,\mathrm{d}t + \left((\mathrm{d}t)^2\right). \tag{8.3}$$

From the composition rule for time evolution, we then have

$$\begin{aligned} U(t+\mathrm{d}t,t_0) &= U(t+\mathrm{d}t,t)U(t,t_0) \\ &= \left(1 - \frac{i}{\hbar}H(t)\,\mathrm{d}t\right)U(t,t_0)\,. \end{aligned} \tag{8.4}$$

We can rearrange this to give

$$U(t+\mathrm{d}t,t_0) - U(t,t_0) = -\frac{i}{\hbar}H(t)\,\mathrm{d}tU(t,t_0)\,. \tag{8.5}$$

Dividing both sides by $\mathrm{d}t/i\hbar$ and taking the limit $\mathrm{d}t \to 0$, this becomes

$$i\hbar\frac{\mathrm{d}U(t,t_0)}{\mathrm{d}t} = H(t)U(t,t_0)\,. \tag{8.6}$$

We now discuss solutions to Eq. (8.6) in several cases:

1. If $H(t) = H$, this result reduces to our previous result,

$$U(t,t_0) = e^{-iH(t-t_0)/\hbar}\,. \tag{8.7}$$

2. If $[H(t), H(t')] = 0$ for all t, t', then we can simultaneously diagonalize the Hamiltonian at all times, meaning we can choose a basis of states that are eigenstates of $H(t)$ for all time (the associated eigenvalues may change as a function of time). We then have

$$U(t,t_0) = \exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\,H\big(t'\big)\right). \tag{8.8}$$

We can check that this is the correct expression by going to the diagonal basis of $H(t)$ and considering the action of $U(t,t_0)$ in this basis, or by differentiating the right-hand side to see that it satisfies Eq. (8.6).

3. In the most general case, $[H(t), H(t')] \neq 0$. An example is a spin-$\frac{1}{2}$ particle in a magnetic field that changes orientation as a function of time. What can we say in this case?

   We know that the Hamiltonian completely defines the time evolution for an infinitesimal time step, and we know that we can build up finite time evolution from infinitesimal time evolution using the composition law. We discretize time into N steps,

   $$t_0 < t_1 < t_2 < \cdots < t_{N-1} < t\,, \tag{8.9}$$

   with

   $$\Delta t = t_{i+1} - t_i = \frac{t - t_0}{N}\,. \tag{8.10}$$

   We could choose each time interval to have a different length, but the ultimate result will not be affected by our choice.

   We take N large, so that Δt is small. Then,

   $$U(t_{i+1}, t_i) = 1 - \frac{i}{\hbar} H(t_i)\Delta t + O\big((\Delta t)^2\big)\,. \tag{8.11}$$

   If we are ignoring terms of order $(\Delta t)^2$, then we can write

   $$U(t_{i+1}, t_i) \approx e^{-iH(t_i)\Delta t/\hbar}\,. \tag{8.12}$$

   We then build up the finite time-evolution operator as

   $$U(t, t_0) = U(t, t_{N-1})U(t_{N-1}, t_{N-2})\cdots U(t_{i+1}, t_i)\cdots U(t_1, t_0) = \prod_{i=0}^{N-1} U(t_{i+1}, t_i)\,. \tag{8.13}$$

   Note that the ordering of these operators is crucial, because these operators do not commute. We must ensure that the operators are ordered so that later times are to the left. This prescription for ordering of the operators is called *time ordering*, and the right-hand side of Eq. (8.13) is called a *time-ordered product*.

   We want to take the limit $N \to \infty, \Delta t \to 0$ with $N\Delta t$ held fixed. Note that every term in the product Eq. (8.13) is an exponential, and so we may be tempted to write the product as the exponential of a sum of operators. We cannot do this, strictly speaking, because the operators do not commute. However, we can invent notation, and formally write

   $$U(t, t_0) = \mathrm{T}\left[\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t'\; H\big(t'\big)\right)\right]. \tag{8.14}$$

   This is called a *time-ordered exponential*, and is defined to be the product in Eq. (8.13) in the limit $N \to \infty, \Delta t \to 0$ with $N\Delta t$ held constant. The operator T is called the *time-ordering operator*; it reorders the operators in its argument so that they are time-ordered (disregarding commutation rules when moving the operators around). What guarantees that this limit is well-defined and exists? This is guaranteed because the operator $U(t, t_0)$ is well-defined, and can be written as a composition of infinitesimal time evolution operators for any partition of the time interval $[t_0, t]$.

   An alternate approach is to solve Eq. (8.6) as a formal power series. We can integrate Eq. (8.6) to reach

   $$\int_{t_0}^{t} \mathrm{d}t'\; \frac{\mathrm{d}}{\mathrm{d}t}\big(U\big(t', t_0\big)\big) = -\frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t'\; H\big(t'\big)U\big(t', t_0\big)\,. \tag{8.15}$$

If we carry out the integration on the left-hand side, we find

$$\int_{t_0}^{t} \mathrm{d}t' \, \frac{\mathrm{d}}{\mathrm{d}t}\big(U(t',t_0)\big) = U(t,t_0) - U(t_0,t_0) = U(t,t_0) - 1\,. \tag{8.16}$$

Subtracting the 1 to the right-hand side of Eq. (8.15), we have

$$U(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t' \, H(t')U(t',t_0)\,. \tag{8.17}$$

This is an expression in terms of the Hamiltonian and time-evolution operator at times t' with $t' \leq t$. We can similarly write

$$U(t',t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t'} \mathrm{d}t'' \, H(t'')U(t'',t_0)\,, \tag{8.18}$$

where the integrand is evaluated at values $t'' \leq t'$. We can carry this process out an infinite number of times and compose the results to give

$$\begin{aligned} U(t,t_0) = 1 &- \frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t' \, H(t') + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} H(t')H(t'') + \cdots \\ &+ \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t_1} \mathrm{d}t_2 \cdots \int_{t_0}^{t_{n-1}} \mathrm{d}t_n \, H(t_1)\cdots H(t_n) + \cdots\,. \end{aligned} \tag{8.19}$$

Note that the operators in each integrand are time-ordered. This observation allows us to write Eq. (8.19) in such a way that the limits of integration are not so complicated, using

$$\int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t'' \, H(t')H(t'') = \frac{1}{2}\int_{t_0}^{t} \int_{t_0}^{t} \mathrm{T}\big[H(t')H(t'')\big]\,. \tag{8.20}$$

The factor of $\frac{1}{2}$ here deals with overcounting. If we rewrite each term in Eq. (8.19) in a similar way, there will be a factor of $\frac{1}{n!}$ on the nth term to deal with overcounting. We then have

$$\begin{aligned} U(t,t_0) = 1 &- \frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t_1 \, H(t_1) + \cdots \\ &+ \frac{1}{n!}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} \mathrm{d}t_1 \int_{t_0}^{t} \mathrm{d}t_2 \cdots \int_{t_0}^{t} \mathrm{d}t_n \, \mathrm{T}[H(t_1)\cdots H(t_n)] + \cdots\,. \end{aligned} \tag{8.21}$$

This series is known as the *Dyson series*. This looks like the series expansion of an exponential, with each term time-ordered. Thus, we have reproduced Eq. (8.14) with a different approach,

$$U(t,t_0) = \mathrm{T}\left[\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t' \, H(t')\right)\right]. \tag{8.22}$$

8.2 Interaction Picture

We have seen the Schrödinger picture and the Heisenberg picture; now we will discuss a third picture, due to Dirac, called the *interaction picture*. This is a mixed picture that is useful when we can write the Hamiltonian in the form

$$H(t) = H_0(t) + V(t)\,, \tag{8.23}$$

where we have a very good understanding of the dynamics under the Hamiltonian $H_0(t)$, and we think of $V(t)$ as a small perturbation to the system described by $H_0(t)$.

In the interaction picture, we remove the evolution due to H_0 from the state by writing

$$|\psi_{\rm I}(t)\rangle = U_0^{-1}(t)|\psi_{\rm S}(t)\rangle\,, \tag{8.24}$$

with $U_0(t)$ the time-evolution operator generated by H_0. Contrast this with the expression in the Heisenberg picture, where the states were defined as

$$|\psi_{\rm H}(t)\rangle = U^{-1}(t)|\psi_{\rm S}(t)\rangle\,, \tag{8.25}$$

The state $|\psi_{\rm I}(t)\rangle$ evolves in time according to

$$i\hbar\frac{\rm d}{{\rm d}t}|\psi_{\rm I}(t)\rangle = V_{\rm I}(t)|\psi_{\rm I}(t)\rangle\,, \tag{8.26}$$

where

$$V_{\rm I}(t) = U_0^{-1}(t)V(t)U_0(t)\,. \tag{8.27}$$

Operators in the interaction picture evolve according to

$$A_{\rm I}(t) = U_0^{-1}(t)AU_0(t)\,. \tag{8.28}$$

The interaction picture is useful when we fully understand the dynamics of H_0, and V is a weak probe. We can then use this approach to determine how quickly our system absorbs energy from the weak probe. In the interaction picture, we are hiding the time evolution due to the system we understand, so that we only see the time evolution coming from the unknown part of the system.

We can now derive the equation of motion for the time-evolution operator in the interaction picture. We define

$$U(t) := U_0(t)U_{\rm I}(t)\,, \tag{8.29}$$

where $U_{\rm I}(t)$ captures the time-evolution due to the perturbation. We know that the time-evolution operator must satisfy

$$i\hbar\frac{\rm d}{{\rm d}t}U(t) = (H_0+V)U(t)\,. \tag{8.30}$$

By definition, $U_0(t)$ is the time-evolution operator in the system with Hamiltonian $H_0(t)$, so it satisfies

$$i\hbar\frac{\rm d}{{\rm d}t}U_0(t) = H_0U_0(t)\,. \tag{8.31}$$

We compute

$$i\hbar\frac{\rm d}{{\rm d}t}U(t) = i\hbar\frac{\rm d}{{\rm d}t}(U_0U_{\rm I}) = i\hbar\left(\frac{{\rm d}U_0}{{\rm d}t}U_{\rm I} + U_0\frac{{\rm d}U_{\rm I}}{{\rm d}t}\right). \tag{8.32}$$

Using Eq. (8.31) on the first term of the right-hand side yields

$$i\hbar\frac{\rm d}{{\rm d}t}U(t) = H_0U_0U_{\rm I} + i\hbar U_0\frac{{\rm d}U_{\rm I}}{{\rm d}t} = H_0U + i\hbar U_0\frac{{\rm d}U_{\rm I}}{{\rm d}t}\,. \tag{8.33}$$

On the other hand, from Eq. (8.30), we have

$$i\hbar\frac{\rm d}{{\rm d}t}U(t) = H_0U + VU\,. \tag{8.34}$$

Thus,

$$i\hbar U_0\frac{{\rm d}U_{\rm I}}{{\rm d}t} = VU = VU_0U_{\rm I}\,, \tag{8.35}$$

which gives us

$$i\hbar\frac{{\rm d}U_{\rm I}}{{\rm d}t} = U_0^{-1}VU_0U_{\rm I} = V_{\rm I}(t)U_{\rm I}(t)\,. \tag{8.36}$$

Lecture 9 (Oct. 4, 2017)

9.1 Spin-$\frac{1}{2}$ in an AC Field

Consider a spin-$\frac{1}{2}$ system in a time-dependent magnetic field. The Hamiltonian is

$$H = -\frac{ge}{2m}\boldsymbol{S}\cdot\boldsymbol{B}(t)\,. \tag{9.1}$$

We will consider a particular class of time-dependent magnetic fields, which we can write in the form

$$\boldsymbol{B}(t) = V_0\hat{\boldsymbol{z}} + B_1(\cos(\omega t)\hat{\boldsymbol{x}} + \sin(\omega t)\hat{\boldsymbol{y}})\,. \tag{9.2}$$

We have previously studied the case where $B_1 = 0$.

In the case $B_1 = 0$, there are two energy eigenstates — the spin-up and spin-down states along the z-axis — that have an energy splitting of

$$\Delta E = \left|\frac{geB_0\hbar}{2m}\right|. \tag{9.3}$$

Now that we are considering a field with nonzero B_1, we expect the system to be able to absorb energy from the oscillating component of the field to transition between the two energy eigenstates of the static field.

We will analyze this problem in the interaction picture. A seemingly sensible choice is to regard the B_0 term of the Hamiltonian to be H_0, the part of the Hamiltonian for which the resulting time evolution is already understood. We write

$$H = H_0 + V(t)\,, \tag{9.4}$$

where

$$\begin{aligned} H_0 &= -\frac{ge}{2m}S^z B\,,\\ V(t) &= -\frac{ge}{2m}\boldsymbol{S}_\perp\cdot\boldsymbol{B}_\perp(t)\,. \end{aligned} \tag{9.5}$$

Here, we have defined

$$\begin{aligned} \boldsymbol{S}_\perp &:= S^x\hat{\boldsymbol{x}} + S^y\hat{\boldsymbol{y}}\,,\\ \boldsymbol{B}_\perp &:= B_1(\cos(\omega t)\hat{\boldsymbol{x}} + \sin(\omega t)\hat{\boldsymbol{y}})\,. \end{aligned} \tag{9.6}$$

We define

$$\omega_0 = \left|\frac{geB_0}{2m}\right|, \tag{9.7}$$

so that we have

$$H_0 = \omega_0 S_z\,. \tag{9.8}$$

(Note that we are taking $e < 0$ so that the signs work out here.) The time-evolution operator due to this part of the Hamiltonian is

$$U_0(t) = e^{-iH_0 t/\hbar}\,. \tag{9.9}$$

The equation of motion in the interaction picture is

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi_{\mathrm{I}}(t)\rangle = V_{\mathrm{I}}(t)|\psi_{\mathrm{I}}(t)\rangle\,, \tag{9.10}$$

where

$$\begin{aligned} V_{\rm I}(t) &= U_0^\dagger(t) V(t) U_0(t) \\ &= e^{i\omega_0 S^z t/\hbar} V(t) e^{-i\omega_0 S^z t/\hbar} \\ &= -\frac{geB_1}{2m} e^{i\omega_0 S^z t/\hbar} (S^x \cos(\omega t) + S^y \sin(\omega t)) e^{-i\omega_0 S^z t/\hbar} \,. \end{aligned} \tag{9.11}$$

Before evaluating this directly, it is convenient to define

$$S^\pm = S^x \pm i S^y \,. \tag{9.12}$$

These are called the *ladder operators*. Note that they are Hermitian conjugates of one another. We can now compute

$$\begin{aligned} e^{i\phi\sigma^z/2} S^+ e^{-i\phi\sigma^z/2} &= \left(\cos\frac{\phi}{2} + i\sigma^z \sin\frac{\phi}{2}\right) S^+ \left(\cos\frac{\phi}{2} - i\sigma^z \sin\frac{\phi}{2}\right) \\ &= S^+ \left(\cos\frac{\phi}{2} - i\sigma^z \sin\frac{\phi}{2}\right)^2 \\ &= S^+ (\cos\phi - i\sigma^z \sin\phi) \\ &= S^+ e^{i\phi} \,. \end{aligned} \tag{9.13}$$

Here, in the first line we have expanded the exponentials, using the fact that the Pauli matrices square to the identity. In the second line, we have used the fact that σ^z anticommutes with both σ^x and σ^y. Note that

$$\sigma^x + i\sigma^y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + i \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} . \tag{9.14}$$

This is why these operators are called ladder operators (S^+ specifically is the *raising operator*). Taking the Hermitian conjugate of Eq. (9.13), we have

$$e^{i\phi\sigma^z/2} S^- e^{-i\phi\sigma^z/2} = S^- e^{-i\phi} \,. \tag{9.15}$$

Now we can return to the calculation of $V_{\rm I}(t)$. We have

$$\begin{aligned} V_{\rm I}(t) &= -\frac{geB_1}{2m} e^{i\omega_0 S^z t/\hbar} \left(\frac{S^+ e^{i\omega t} + S^- e^{-i\omega t}}{2}\right) e^{-i\omega_0 S^z t/\hbar} \\ &= -\frac{geB_1}{4m} \left(S^+ e^{i(\omega_0+\omega)t} + S^- e^{-i(\omega_0+\omega)t}\right) . \end{aligned} \tag{9.16}$$

9.1.1 Resonant Drive

First, let's consider the simple case where the external frequency is $\omega = -\omega_0$. This case is called "resonant drive." In this case, the potential in Eq. (9.16) simplifies to

$$V_{\rm I} = -\frac{geB_1}{4m} \left(S^+ + S^-\right) = -\frac{geB_1}{4m} S^x \,, \tag{9.17}$$

which is time-independent. The interaction picture time-evolution operator is then

$$U_{\rm I}(t) = e^{i\frac{geB_1}{2m\hbar} S^x t} \,. \tag{9.18}$$

This operator rotates states about the x-axis of spin space by an angle $\frac{geB_1}{2m\hbar} t$. The frequency of rotation,

$$\omega_{\rm R} = \left| \frac{geB_1}{2m} \right| , \tag{9.19}$$

is called the *Rabi frequency*. Note that the time-evolution operator,

$$U_{\mathrm{I}}(t) = e^{-i\omega_{\mathrm{R}} S^x t/\hbar} = e^{-i\left(\frac{\omega_{\mathrm{R}} t}{2}\right)\sigma^x}\,, \tag{9.20}$$

does not oscillate with a frequency of $\frac{2\pi}{\omega_{\mathrm{R}}}$, but rather $\frac{4\pi}{\omega_{\mathrm{R}}}$,

$$U_{\mathrm{I}}\left(t + \frac{4\pi}{\omega_{\mathrm{R}}}\right) = U_{\mathrm{I}}(t)\,. \tag{9.21}$$

This operator has periodicity twice that of any observables in the system. What happens after a time $\frac{2\pi}{\omega_{\mathrm{R}}}$? At this time, we have

$$U_{\mathrm{I}}\left(\frac{2\pi}{\omega_{\mathrm{R}}}\right) = e^{-i\pi\sigma^x} = -\mathbb{1}\,. \tag{9.22}$$

Thus, any observables will oscillate at the Rabi frequency, but if we were able to measure the phase of a state, we would find that it oscillates at half the Rabi frequency.

Imagine that we start in a spin-up state, i.e.,

$$|\psi_{\mathrm{I}}(0)\rangle \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{9.23}$$

where the arrow indicates that we are representing the state as a vector. At time t, we then have

$$\begin{aligned} |\psi_{\mathrm{I}}(t)\rangle &\to e^{-i\omega_{\mathrm{R}} t\sigma^x/2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \left(\cos\left(\frac{\omega_{\mathrm{R}} t}{2}\right) - i\sigma^x \sin\left(\frac{\omega_{\mathrm{R}} t}{2}\right)\right)\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\left(\frac{\omega_{\mathrm{R}} t}{2}\right) & -i\sin\left(\frac{\omega_{\mathrm{R}} t}{2}\right) \\ -i\sin\left(\frac{\omega_{\mathrm{R}} t}{2}\right) & \cos\left(\frac{\omega_{\mathrm{R}} t}{2}\right) \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\left(\frac{\omega_{\mathrm{R}} t}{2}\right) \\ -i\sin\left(\frac{\omega_{\mathrm{R}} t}{2}\right) \end{pmatrix}. \end{aligned} \tag{9.24}$$

Thus, after a time $t = \frac{\pi}{\omega_{\mathrm{R}}}$, we have

$$|\psi_{\mathrm{I}}(\pi/\omega_{\mathrm{R}})\rangle \to \begin{pmatrix} 0 \\ -i \end{pmatrix}, \tag{9.25}$$

and after a time $t = \frac{2\pi}{\omega_{\mathrm{R}}}$, we have

$$|\psi_{\mathrm{I}}(2\pi/\omega_{\mathrm{R}})\rangle \to \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \tag{9.26}$$

where we see that the state has picked up a phase of π. We find that the state oscillates; this behavior is referred to as *Rabi oscillation*.

9.1.2 Off-Resonant Drive

Up until now, we have been working in the case of resonant drive. What happens if we're off resonance, i.e., $\omega \neq -\omega_0$? In this case, even though we went to the interaction picture, we still have a time-dependent Hamiltonian that we don't know how to solve, because $V_{\mathrm{I}}(t)$ is still time-dependent. It seems like we have gained nothing.

The issue is that we made a particular choice for which part of the Hamiltonian H would be H_0, but this was not actually the most convenient choice. Let's instead choose

$$H = \tilde{H}_0 + \tilde{V}(t)\,, \tag{9.27}$$

where

$$\begin{aligned} \tilde{H}_0 &= \omega S^z\,, \\ \tilde{V}(t) &= -\gamma(B_{\text{eff}} S^z + \boldsymbol{B}_\perp(t)\cdot\boldsymbol{S}_\perp)\,. \end{aligned} \tag{9.28}$$

Here, we have defined

$$\begin{aligned} B_{\text{eff}} &:= B_0 + \frac{\omega}{\gamma}\,, \\ \gamma &:= \frac{ge}{2m}\,. \end{aligned} \tag{9.29}$$

Now, let's work in the interaction picture with these choices. Going to the interaction picture is going to a rotating frame, that rotates with $\tilde{H}_0$. Previously, we went to a frame that rotated about the z-axis at frequency ω_0; now we are going to a frame that rotates about the z-axis at the external frequency ω.

We compute

$$\begin{aligned} \tilde{V}_{\text{I}}(t) &= \tilde{U}_0^\dagger(t)\tilde{V}(t)\tilde{U}_0(t) \\ &= e^{i\omega S^z t/\hbar}\left[-\gamma B_{\text{eff}} S^z - \frac{\gamma B_1}{2}\left(S^+ e^{i\omega t} + S^- e^{-i\omega t}\right)\right] e^{-i\omega S^z t/\hbar}\,. \end{aligned} \tag{9.30}$$

The first term in brackets is unaffected by the time evolution, because S^z commutes with the exponentials outside the brackets. The second term in brackets, along with the time-evolution exponentials, looks exactly like the calculation we did in Eq. (9.16), but with the frequencies in all of the exponentials exactly matched. Thus, we find

$$\tilde{V}_{\text{I}}(t) = -\gamma B_{\text{eff}} S^z - \gamma B_1 S^x\,. \tag{9.31}$$

This is time-independent, and so we now declare the problem solved. The spin precesses about the direction of

$$\tilde{\boldsymbol{B}} = (B_1, 0, B_{\text{eff}})\,. \tag{9.32}$$

These are modified Rabi oscillations, with a frequency of

$$\omega_{\text{R}} = \gamma\sqrt{B_{\text{eff}}^2 + B_1^2}\,. \tag{9.33}$$

9.2 Path Integral Formulation of Quantum Mechanics

Time evolution in the Schrödinger picture is given by

$$|\psi(t)\rangle = \sum_{a'} c_{a'}(t)\big|a'\big\rangle\,, \tag{9.34}$$

where the $|a'\rangle$ form an energy eigenbasis, and

$$c_{a'}(t) = e^{-iE_{a'}(t-t_0)/\hbar} c_{a'}(t_0)\,. \tag{9.35}$$

We can consider these particles in position space, and define

$$u_{a'}(\boldsymbol{x}) = \big\langle \boldsymbol{x}\big|a'\big\rangle \tag{9.36}$$

to be the energy eigenfunctions. Then the position-space wavefunction is given by

$$\psi(\boldsymbol{x},t)=\sum_{a'}e^{-iE_{a'}(t-t_0)/\hbar}c_{a'}(t_0)u_{a'}(\boldsymbol{x})\,. \tag{9.37}$$

We can rewrite this as

$$\psi(\boldsymbol{x},t)=\int_{-\infty}^{\infty}\mathrm{d}^dx'\;K\big(\boldsymbol{x},t;\boldsymbol{x}',t_0\big)\psi\big(\boldsymbol{x}',t_0\big)\,, \tag{9.38}$$

with

$$K\big(\boldsymbol{x},t;\boldsymbol{x}',t_0\big)=\sum_{a'}\big\langle\boldsymbol{x}\big|a'\big\rangle e^{-iE_{a'}(t-t_0)/\hbar}\big\langle a'\big|\boldsymbol{x}'\big\rangle=\big\langle\boldsymbol{x}\big|U(t,t_0)\big|\boldsymbol{x}'\big\rangle\,. \tag{9.39}$$

The object $K(\boldsymbol{x},t;\boldsymbol{x}',t_0)$ is called the *propagator*.

Lecture 10 (Oct. 11, 2017)

10.1 Path Integral Formulation of Quantum Mechanics

10.1.1 The Propagator

In the last lecture, we introduced the propagator,

$$K(\boldsymbol{x},t;\boldsymbol{x}',t_0)=\sum_{a'}\langle\boldsymbol{x}|a'\rangle e^{-iE_{a'}(t-t_0)/\hbar}\langle a'|\boldsymbol{x}'\rangle=\langle\boldsymbol{x}|U(t,t_0)|\boldsymbol{x}'\rangle\,. \tag{10.1}$$

In terms of the propagator, we can write the wavefunction in the form

$$\psi(\boldsymbol{x},t)=\int_{-\infty}^{\infty}\mathrm{d}^dx'\;K(\boldsymbol{x},t;\boldsymbol{x}',t_0)\psi(\boldsymbol{x}',t_0)\,, \tag{10.2}$$

From the Schrödinger equation, assuming a Hamiltonian of the form

$$H=\frac{\boldsymbol{p}^2}{2m}+V(\boldsymbol{x})\,, \tag{10.3}$$

the propagator must satisfy the differential equation

$$\left(i\hbar\partial_t+\frac{\hbar^2}{2m}\nabla^2-V(\boldsymbol{x})\right)K(\boldsymbol{x},t;\boldsymbol{x}',t_0)=0\,. \tag{10.4}$$

Furthermore, it must satisfy

$$\lim_{t\to t_0}K(\boldsymbol{x},t;\boldsymbol{x}',t_0)=\delta^{(d)}(\boldsymbol{x}-\boldsymbol{x}')\,. \tag{10.5}$$

A convenient related quantity is the *retarded propagator*

$$K_{\mathrm{ret}}(\boldsymbol{x},t;\boldsymbol{x}',t_0)=\theta(t-t_0)K(\boldsymbol{x},t;\boldsymbol{x}',t_0)\,. \tag{10.6}$$

The retarded propagator satisfies

$$\left(i\hbar\partial_t+\frac{\hbar^2}{2m}\nabla^2-V(\boldsymbol{x})\right)K_{\mathrm{ret}}(\boldsymbol{x},t;\boldsymbol{x}',t')=i\hbar\delta^{(d)}(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t')\,, \tag{10.7}$$

i.e., it is a Green's function for the Schrödinger equation (a solution of the Schrödinger equation with a delta function source or forcing term).

Alternatively, we can think of the propagator as

$$K(\boldsymbol{x},t;\boldsymbol{x}',t_0)=\langle\boldsymbol{x},t|\boldsymbol{x}',t_0\rangle\,, \tag{10.8}$$

which follows from its definition.

As an example, consider the free particle in one dimension, with Hamiltonian

$$H=\frac{p^2}{2m}\,. \tag{10.9}$$

The propagator in this case is given by

$$\begin{aligned}K(x,t;x',t')&=\int\mathrm{d}p\;\langle x|p\rangle e^{-\frac{i}{\hbar}\frac{p^2}{2m}(t-t')}\langle p|x'\rangle\\&=\int\frac{\mathrm{d}p}{2\pi\hbar}\;e^{ip(x-x')-\frac{i}{\hbar}\frac{p^2}{2m}(t-t')}\\&=\sqrt{\frac{m}{2\pi i\hbar(t-t')}}e^{i\frac{m(x-x')^2}{2\hbar(t-t')}}\,.\end{aligned} \tag{10.10}$$

The propagator has a number of properties that make it useful to calculate. The first is related to quantum statistical mechanics. We define the partition function as

$$G(t) = \int \mathrm{d}^d x \; K(\boldsymbol{x}, t; \boldsymbol{x}, 0) = \sum_{a'} e^{-iE_{a'}t/\hbar} \,. \tag{10.11}$$

If we analytically continue to imaginary time, $t = -i\hbar\beta$, then this yields the familiar partition function of statistical mechanics:

$$G(-i\hbar\beta) = \sum_{a'} e^{-\beta E_{a'}} := Z \,. \tag{10.12}$$

Another useful quantity is the Fourier transform of the partition function,

$$\tilde{G}(E) = -i \int \mathrm{d}t \; G_{\mathrm{ret}}(t) e^{iEt/\hbar} = -i \sum_{a'} \int_0^\infty \mathrm{d}t \; e^{i(E-E_{a'})t/\hbar} \,, \tag{10.13}$$

where $G_{\mathrm{ret}}(t) = \theta(t)G(t)$. In order to make this integral converge, we take $E \to E + i\epsilon$, which then yields

$$\tilde{G}(E) = \sum_{a'} \frac{\hbar}{E - E_{a'} + i\epsilon} \,. \tag{10.14}$$

The poles of $\tilde{G}$ in the limit $\epsilon \to 0$ describe the energy spectrum of the system.

We can also use the Fourier transform of the partition function to find the density of states. The density of states is defined as

$$\rho(E) = \sum_{a'} \delta(E - E_{a'}) \,. \tag{10.15}$$

Using the property

$$\begin{aligned} \lim_{\epsilon\to 0} \mathrm{Im} \, \frac{1}{E - E' + i\epsilon} &= \lim_{\epsilon \to 0} \frac{-\epsilon}{(E-E')^2 + \epsilon^2} \\ &= -\pi \delta\big(E - E'\big) \,, \end{aligned} \tag{10.16}$$

we see that we can write the density of states as

$$\rho(E) = -\frac{1}{\pi\hbar} \, \mathrm{Im} \, \tilde{G}(E) \,. \tag{10.17}$$

10.1.2 Path Integrals

From the composition property of time evolution,

$$U(t, t_0) = U\big(t, t'\big) U\big(t', t_0\big) \,, \quad t > t' > t_0 \,, \tag{10.18}$$

we see that the propagator satisfies a similar composition property,

$$K\big(\boldsymbol{x}, t; \boldsymbol{x}', t_0\big) = \int \mathrm{d}^d \tilde{x} \; K\big(\boldsymbol{x}, t; \tilde{\boldsymbol{x}}, \tilde{t}\big) K\big(\tilde{\boldsymbol{x}}, \tilde{t}; \boldsymbol{x}', t_0\big) \,. \tag{10.19}$$

We can break the interval $[t_0, t]$ into N equal time intervals of width

$$\Delta t = \frac{t - t_0}{N} \,, \tag{10.20}$$

with endpoints

$$t_0 < t_1 < \cdots < t_N = t\,. \tag{10.21}$$

Iterating the composition rule then gives us

$$K(x_N, t_N; x_0, t_0) = \int \prod_{k=1}^{N-1} K(x_N, t_N; x_{N-1}, t_{N-1}) \cdots K(x_2, t_2; x_1, t_1) K(x_1, t_1; x_0, t_0)\,. \tag{10.22}$$

Feynman proposed that

$$K\big(x, t; x', t_0'\big) = \int [\mathcal{D}x] e^{iS[x(t)]/\hbar}\,, \tag{10.23}$$

where the right-hand side is a sum over all possible paths from (x', t_0) to (x, t) (including those that do not satisfy the equations of motion), and S is the classical action of the trajectory. Recall that in classical mechanics, the Lagrangian is given by

$$L = \frac{1}{2} m\dot{x}^2 - V(x)\,, \tag{10.24}$$

and the action is given by $S = \int \mathrm{d}t\; L$. (Here, we have assumed that the kinetic energy is of the form $m\dot{x}^2/2$.) The classical trajectory between two spacetime points is obtained by extremizing this action, $\frac{\delta S}{\delta x} = 0$. This is known as the principle of least action. This requirement leads to the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0\,. \tag{10.25}$$

For the Lagrangian given above, this gives Newton's force law,

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\frac{\partial V}{\partial x}\,. \tag{10.26}$$

Note that Feynman's proposed expression for the propagator,

$$K\big(x, t; x', t_0'\big) = \int [\mathcal{D}x] e^{iS[x(t)]/\hbar}\,, \tag{10.27}$$

clearly satisfies the composition property. It also gives a very simple connection to classical mechanics: in the $\hbar \to 0$ limit, we expect the sum over paths is dominated by the path for which the phase is stationary (this is the stationary phase approximation, which we will discuss more thoroughly later), i.e., the path for which S is extremized. Thus, Feynman's path integral expression yields the principle of least action in the classical limit.

Lecture 11 (Oct. 16, 2017)

11.1 Path Integrals

As we have seen, the amplitude for a state to "propagate" from coordinates (x,t) to (x',t') is given by the propagator,

$$K\big(x',t';x,t\big) = \int [\mathcal{D}x]\; e^{iS[x(t)]/\hbar}\,, \tag{11.1}$$

where $S[x(t)]$ is the classical action of the path $x(t)$,

$$S[x(t)] = \int \mathrm{d}t'\; L\big(x,\dot{x},t'\big)\,, \tag{11.2}$$

for L the classical Lagrangian. The integration in the path integral is over all possible paths $x(t)$.

We need to define this precisely. In particular, we need to write a measure on the space of paths in order to know how to properly integrate over all paths. The strategy will be to start with the definition of K, and then derive the path integral and the appropriate definition of the measure. Then we will go back and recompute K for a free particle.

The starting point of deriving the path integral is to use the composition law for K, which comes from the composition law for the time-evolution operator. Recall that the propagator K is the matrix element

$$K\big(x,t;x',t'\big) = \langle x|U\big(t',t\big)|x\rangle\,. \tag{11.3}$$

From the composition law, we have

$$\begin{aligned} K(x_N,t_N;x_0,t_0) &= \int \prod_{k=1}^{N-1} \mathrm{d}x_k\; K(x_N,t_N;x_{N-1},t_{N-1})\cdots K(x_1,t_1;x_0,t_0) \\ &= \int \prod_{k=1}^{N-1} \mathrm{d}x_k\; \langle x_N|U(t_N,t_{N-1})|x_{N-1}\rangle\cdots\langle x_1|U(t_1,t_0)|x_0\rangle \\ &= \int \prod_{k=1}^{N-1} \mathrm{d}x_k\; \langle x_N|e^{-i\epsilon H/\hbar}|x_{N-1}\rangle\cdots\langle x_1|e^{-i\epsilon H/\hbar}|x_0\rangle\,, \end{aligned} \tag{11.4}$$

where $\epsilon = t/N$. This expression holds in general for time-dependent Hamiltonians, if we understand each occurrence of H in the expression above to be a function of the time.

We now consider a single matrix element in this expression,

$$\langle x_{j+1}|e^{-i\epsilon H/\hbar}|x_j\rangle = \langle x_{j+1}|e^{-i\epsilon(T+V)/\hbar}|x_j\rangle\,, \tag{11.5}$$

where

$$T = \frac{p^2}{2m}\,, \quad V = V(x)\,. \tag{11.6}$$

Note that

$$e^{-i\epsilon(T+V)/\hbar} = 1 - \frac{i\epsilon(T+V)}{\hbar} - \frac{\epsilon^2(T+V)^2}{2\hbar^2} + \cdots = e^{-i\epsilon T/\hbar}e^{-i\epsilon V/\hbar} + O\big(\epsilon^2\big)\,. \tag{11.7}$$

Thus, if we take $\epsilon\to 0, N\to\infty$ with $N\epsilon$ fixed, then we can use

$$e^{-i\epsilon(T+V)/\hbar} \xrightarrow{\epsilon\to 0} e^{-i\epsilon T/\hbar}e^{-i\epsilon V/\hbar}\,. \tag{11.8}$$

We can only make this replacement legally in this limit, as these two expressions differ at higher orders.

We now choose ϵ small so that we can safely omit the $O(\epsilon^2)$ terms above. We then have

$$\begin{aligned}\langle x_{j+1}|e^{-i\epsilon H/\hbar}|x_j\rangle &\xrightarrow{\epsilon\to 0} \langle x_{j+1}|e^{-i\epsilon p^2/2m\hbar}e^{-i\epsilon V/\hbar}|x_j\rangle \\ &= \langle x_{j+1}|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle e^{-i\epsilon V(x_j)/\hbar}\,.\end{aligned} \tag{11.9}$$

We see then that we only have to evaluate the matrix element for a free particle. We can evaluate this by inserting the resolution of the identity in the momentum basis, which results in taking a Fourier transform:

$$\begin{aligned}\langle x_{j+1}|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle &= \int \mathrm{d}p_j\ \langle x_{j+1}|p_j\rangle\langle p_j|e^{-i\epsilon p^2/2m\hbar}|x_j\rangle \\ &= \int \frac{\mathrm{d}p_j}{2\pi\hbar}e^{ip_j(x_{j+1}-x_j)/\hbar}e^{-i\epsilon p_j^2/2m\hbar} \\ &= \sqrt{\frac{m}{2\pi i\hbar\epsilon}}e^{\frac{im(x_{j+1}-x_j)^2}{2\hbar\epsilon}}\,.\end{aligned} \tag{11.10}$$

This is the same result we found in the last lecture.

Thus, we have

$$\langle x_{j+1}|e^{-i\epsilon H/\hbar}|x_j\rangle \xrightarrow{\epsilon\to 0} \sqrt{\frac{m}{2\pi i\hbar\epsilon}}e^{i\left[\frac{m(x_{j+1}-x_j)^2}{2\hbar\epsilon}-\frac{\epsilon V(x_j)}{\hbar}\right]}\,. \tag{11.11}$$

We then have

$$K(x_N,t_N;x_0,t_0) = \lim_{\substack{\epsilon\to 0\\ N\to\infty\\ N\epsilon=t_N-t_0}}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{N/2}\int\prod_{k=1}^{N-1}\mathrm{d}x_k\ e^{i\sum_{j=0}^{N-1}\left[\frac{m(x_{j+1}-x_j)^2}{2\hbar\epsilon}-\frac{\epsilon V(x_j)}{\hbar}\right]}\,. \tag{11.12}$$

In the limit $\epsilon\to 0$, we have

$$x_{j+1}-x_j \approx \epsilon\frac{\mathrm{d}x}{\mathrm{d}t}\,, \tag{11.13}$$

which gives us

$$\frac{(x_{j+1}-x_j)^2}{\epsilon} \approx \epsilon\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2\,. \tag{11.14}$$

Thus,

$$\begin{aligned}\sum_{j=0}^{N-1}\left[\frac{m(x_{j+1}-x_j)^2}{2\hbar\epsilon}-\frac{\epsilon V(x_j)}{\hbar}\right] &\xrightarrow{\epsilon\to 0}\frac{1}{\hbar}\int\mathrm{d}t\ \left[\frac{1}{2}m\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2-V(x)\right] \\ &= \frac{1}{\hbar}\int_{t_0}^{t_N}\mathrm{d}t\ L(x,\dot{x},t)\,,\end{aligned} \tag{11.15}$$

where L is the classical Lagrangian. This is the desired result. We now define the path integral formally through the limiting process

$$\int[\mathcal{D}x]\ e^{iS[x(t)]/\hbar} := \lim_{\substack{\epsilon\to 0\\ N\to\infty\\ N\epsilon=t_N-t_0}}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{N/2}\int\prod_{k=1}^{N-1}\mathrm{d}x_k\ e^{i\sum_{j=0}^{N-1}\left[\frac{m(x_{j+1}-x_j)^2}{2\hbar\epsilon}-\frac{\epsilon V(x_j)}{\hbar}\right]}\,. \tag{11.16}$$

The factor in front of the product of integrals is required to make sure that the limit exists.

Now we will recalculate the free particle propagator. Let us choose the number of steps to be a power of two, $N = 2^n$, and let us define

$$K_N = \left(\frac{mN}{2\pi i\hbar t}\right)^{N/2} \int \prod_{k=0}^{N-1} \mathrm{d}x_k \; e^{\frac{imN}{2\hbar t}\sum_j (x_{j+1}-x_j)^2} \,. \tag{11.17}$$

Again, we have $\epsilon = t/N$. Expanding all of the squares, we see that

$$\sum_j (x_{j+1}-x_j)^2 = x_0^2 + 2x_1^2 + 2x_2^2 + \cdots + 2x_{N-1}^2 + x_N^2 - 2x_0x_1 - 2x_1x_2 - \cdots - 2x_{N-1}x_N \,. \tag{11.18}$$

We now carry out all of the odd integrals first, which will cause the problem to simplify. We find

$$\begin{aligned} K_N &= \int \prod_{j \text{ even}} \mathrm{d}x_j \int \prod_{k \text{ odd}} \mathrm{d}x_k \left(\frac{m2^n}{2\pi i\hbar t}\right) e^{\frac{im2^n}{2\hbar t}\left[2x_k^2 + x_{k-1}^2 + x_{k+1}^2 - 2x_k(x_{k-1}+x_{k+1})\right]} \\ &= \int \prod_{j \text{ even}} \mathrm{d}x_j \int \prod_{k \text{ odd}} \mathrm{d}x_k \left(\frac{m2^n}{2\pi i\hbar t}\right) e^{\frac{im2^n}{2\hbar t}\left[2\left(x_k - \frac{x_{k-1}+x_{k+1}}{2}\right)^2 + \frac{1}{2}\left(x_{k+1}^2 + x_{k-1}^2\right) - x_{k+1}x_{k-1}\right]} \\ &= \left(\frac{m2^{n-1}}{2\pi i\hbar t}\right)^{2^{n-2}} \int \prod_{\ell=0}^{2^{n-1}-1} \mathrm{d}x_{2\ell} \; e^{\frac{im2^{n-1}}{2\hbar t}\sum_{\ell=0}^{2^{n-1}-1}(x_{2\ell+2}-x_{2\ell})^2} \\ &= K_{N/2} \,. \end{aligned} \tag{11.19}$$

We see then that

$$K_{2^n} = K_{2^{n-1}} = K_{2^{n-2}} = \cdots = K_1 = \sqrt{\frac{m}{2\pi i\hbar t}} e^{\frac{im}{2\hbar t}(x_N - x_0)^2} \,. \tag{11.20}$$

Thus, we have recovered the expected result.

We can explicitly evaluate the path integral for any system in which the potential is at most quadratic in $x, \dot{x}$. We will see several examples on the homework.

The path integral formulation seems unwieldy. Why does anyone bother with this? This is a completely different way of formulating quantum mechanics than we have seen previously, but it is completely equivalent to our other formulations, so what is the use? Feynman emphasized that different formulations, even though they are mathematically equivalent, may each be useful for describing different systems. Furthermore, we can use a given formulation to generalize quantum mechanics, and the different formalisms may not be equivalent in these generalizations. Finally, as Feynman says, it feels psychologically different to be thinking about the same physics from a completely different perspective. In practice, path integrals are not usually a practical way of solving problems. However, the path integral gives us quantum intuition, which allows us to say many things about how systems will behave, even if we cannot solve them exactly.

11.1.1 Technical Details

For each time step, we wrote

$$\langle x_{j+1}|e^{-i\epsilon H/\hbar}|x_j\rangle = e^{-\frac{i\epsilon m}{2\hbar}\left(\frac{x_{j+1}-x_j}{2}\right) - \frac{i\epsilon}{\hbar}V(x_j)} \,. \tag{11.21}$$

It seems arbitrary that we wrote

$$e^{-i\epsilon(T+V)/\hbar} \approx e^{-i\epsilon T/\hbar} e^{-i\epsilon V/\hbar} \tag{11.22}$$

instead of

$$e^{-i\epsilon(T+V)/\hbar} \approx e^{-i\epsilon V/\hbar} e^{-i\epsilon T/\hbar}\,, \tag{11.23}$$

in which case we would have had $V(x_{j+1})$ in the exponent of Eq. (11.21) instead of $V(x_j)$. Either is a perfecetly good choice to the order at which we are working. A more symmetric choice is to use $V\left(\frac{x_j+x_{j+1}}{2}\right)$.

This choice is very important for the particle in a magnetic field described by a vector potential $\boldsymbol{A}$ (through $\boldsymbol{B} = \nabla \times \boldsymbol{A}$). In this system, the Lagrangian is

$$L = \frac{1}{2} m \dot{\boldsymbol{x}}^2 + e\boldsymbol{A} \cdot \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} - V(\boldsymbol{x})\,. \tag{11.24}$$

Here, in the quantum mechanical path integral we must use

$$\boldsymbol{A} \to \boldsymbol{A}\left(\frac{x_j + x_{j+1}}{2}\right). \tag{11.25}$$

See the homework for more on this.

Lecture 12 (Oct. 23, 2017)

12.1 Stationary Phase Approximation

Consider an integral of the form

$$\int \mathrm{d}x\, e^{i\lambda f(x)}\,, \tag{12.1}$$

with λ large and $f(x)$ real. Can we use the largeness of λ in order to help evaluate or approximate the integral? Let $f(x)$ have an extremum at $x = x_0$. The claim is that the integral will be dominated by the region around x_0. In most regions of integration, the complex exponential oscillates rapidly as a function of x, and so the integral over any range will sum many phases that cancel out. The region around x_0, however, varies slowly, because $f(x)$ achieves an extremum at x_0. We thus expect that the integral is dominated by x near x_0 (or, in general, near all x_0 for which $f(x)$ achieves an extremum).

More precisely, near x_0, we can write

$$f(x) \approx f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + \cdots\,, \tag{12.2}$$

where the first-order term vanishes because $f(x)$ has an extremum at x_0. We can then write

$$\int \mathrm{d}x\, e^{i\lambda f(x)} \approx e^{i\lambda f(x_0)} \int \mathrm{d}(\delta x)\, e^{i\lambda f''(x_0)(\delta x)^2/2 + \cdots}\,. \tag{12.3}$$

If we assume that δx is small, we can drop the higher order terms in the exponential. However, we can then extend the limits of integration over δx to $\pm\infty$, because the largeness of λ leads to rapid oscillation for large δx, giving cancellation of the exponentials. This yields

$$\begin{aligned} \int \mathrm{d}x\, e^{i\lambda f(x)} &\approx e^{i\lambda f(x_0)} \int_{-\infty}^{\infty} \mathrm{d}(\delta x)\, e^{i\lambda f''(x_0)(\delta x)^2/2} \\ &= e^{i\lambda f(x)} \sqrt{\frac{2\pi i}{\lambda f''(x_0)}} \big(1 + O(\lambda^{-2})\big)\,. \end{aligned} \tag{12.4}$$

We can generalize this to multidimensional integrals, and even path integrals. We thus similarly expect that

$$\int [\mathcal{D}x]\, e^{iS[x(t)]/\hbar} \tag{12.5}$$

is dominated by $x(t) = x_{\mathrm{cl}}(t)$, the classical path, where

$$\left.\frac{\delta S}{\delta x}\right|_{x_{\mathrm{cl}}} = 0\,, \tag{12.6}$$

when $\hbar \to 0$. Then we get

$$K(x_f, t_f; x_i, t_i) \sim e^{iS[x_{\mathrm{cl}}(t)]/\hbar}\,. \tag{12.7}$$

Note that this not only gives us the classical limit, but also gives us interference effects in the near-classical limit: if there are multiple possible classical paths, then in the near-classical limit, these paths can exhibit interference with one another.

12.2 Quantum Particles in Electromagnetic fields

We now discuss the motion of quantum particles in electromagnetic fields. In general, we should assume that every system is quantum mechanical. However, here we will be using a semiclassical approach, in which we will consider a classical background electromagnetic field with quantum particles moving in it. There are many phenomena which require the quantization of the electric field as well as the particles moving in it, but we will not discuss such phenomena here.

12.2.1 Constant Potentials

We start with a potential V. If we change the potential by $V \to V + V_0$, this is a shift in the overall energy, and should have no effect on the particle motion. Classically, this is true because particles move due to forces, not potentials. The force is computed by taking the gradient of the potential, which will not be effected by the shift. In quantum mechanics, the shift in the potential results in an extra phase under time evolution,

$$|\psi(t)\rangle \to e^{-iV_0 t/\hbar}|\psi(t)\rangle\,. \tag{12.8}$$

Thus, the shift serves to change the phase of the wavefunction at each time step. We know that such an overall phase shift is not observable.

However, if we change the potential by a constant shift in one region of space but not in another, then we do expect to see an observable effect. Quantum mechanically, we can interfere particles that have traveled through this region with those that have not.

Imagine that we have a beam of particles being emitted from a source A, and two metallic cages that the beam could pass through, before passing to a detector at B. Assume that the potential is V_1 in the bottom cage and V_2 in the top cage, both constant. The potential is constant within each cage, and so the particles passing through either cage experience no forces. Quantum mechanically, there is nevertheless a difference in the phase factor accumulated while passing through the two different cages:

$$\phi_2 - \phi_1 = \frac{1}{\hbar}\int_{t_i}^{t_f} \mathrm{d}t\,(V_2(t) - V_1(t))\,. \tag{12.9}$$

Let us assume that the two metallic cages are completely identical other than the described potential difference, and that A and B are arranged symmetrically around the two cages. We thus expect, from symmetry, that the amplitude for a particle to travel from A to B would be the same through either cage in absence of the potential difference. If we let a be the amplitude to go from A to B through the bottom cage in time $t_f - t_i$, then the amplitude to go through the top cage must be

$$a e^{-i(\phi_2-\phi_1)}\,. \tag{12.10}$$

The full amplitude to travel from A to B is then

$$a\left(1 + e^{-i(\phi_2-\phi_1)}\right), \tag{12.11}$$

giving a probability of detection of

$$\mathrm{Prob}(A \to B) = 2|a|^2(1 + \cos(\phi_2 - \phi_1))\,. \tag{12.12}$$

Thus, the constant potential difference between the two cages changes physical observables; this is a purely quantum mechanical effect, as the particles experienced no forces during this experiment (neglecting the small forces when entering and exiting the regions of shifted potential). Note that in the limit of $\hbar \to 0$, the phase difference (12.9) oscillates very rapidly. When averaged over any nonzero time interval, these rapidly oscillating phases cancel out, so the interference effects will disappear, as we expect in the classical limit.

12.2.2 Electromagnetic Fields

Let us briefly review electromagnetism. We know that we can represent electric fields $\boldsymbol{E}$ and magnetic fields $\boldsymbol{B}$ through scalar and vector potentials as

$$\boldsymbol{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\,, \quad \boldsymbol{B} = \nabla \times \boldsymbol{A}\,. \tag{12.13}$$

Note that we are using cgs units. The potentials ϕ and $\boldsymbol{A}$ will be extremely useful in describing the motion of quantum mechanical particles through EM fields.

It is useful to introduce a more formal notation, defining the electromagnetic field strength tensor

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu\,, \tag{12.14}$$

with

$$A_\mu = (\phi, \boldsymbol{A})\,, \quad (\mu = 0, 1, 2, 3)\,. \tag{12.15}$$

The electric and magnetic fields are then given by

$$F_{0i} = -F_{i0} = -E_i\,, \quad F_{ij} = \epsilon_{ijk}B_k\,. \tag{12.16}$$

The four-vector potential A_μ is redundant: the gauge transformation

$$A_\mu \to A_\mu + \partial_\mu A \tag{12.17}$$

leaves the field strength $F_{\mu\nu}$ invariant, which means the electric and magnetic fields are unchanged. The field strength tensor is the real physical quantity, and all of the redundant values of A_μ that yield the same field strength tensor describe the same physics.

For a charged particle in an electromagnetic field, the Lagrangian is

$$L = \frac{1}{2}m\dot{\boldsymbol{x}}^2 + \frac{e}{c}\boldsymbol{A}\cdot\dot{\boldsymbol{x}} - e\phi\,. \tag{12.18}$$

In Lagrangian mechanics, which is formulated in terms of position and velocity, there is a notion of momentum as the variable conjugate to the velocity. The *canonical momentum* is given by

$$\boldsymbol{p} := \frac{\partial L}{\partial \dot{\boldsymbol{x}}} = m\dot{\boldsymbol{x}} + \frac{e}{c}\boldsymbol{A}\,. \tag{12.19}$$

Note that the canonical momentum is *not* gauge-invariant; it depends on the choice of gauge.

Using this canonical momentum, we find that the Hamiltonian is given by

$$H = \boldsymbol{p}\cdot\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} - L = \frac{1}{2}m\dot{\boldsymbol{x}}^2 + e\phi\,. \tag{12.20}$$

This is the expected answer; the Hamiltonian is gauge-invariant. Using the definition of the canonical momentum, we can rewrite this as

$$H = \frac{\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2}{2m} + e\phi\,. \tag{12.21}$$

The main conceptual point here is that the *canonical* momentum $\boldsymbol{p}$ is different from $m\boldsymbol{v} = m\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$ (the "kinematic" momentum) when $\boldsymbol{A} \neq 0$. The reason we care about the canonical momentum rather than the kinematic momentum is because it is the quantity that has $\{x, p\} = 1$ (with $\{\cdot,\cdot\}$

the Poisson bracket), which means that it will have the proper commutator with the operator $\boldsymbol{x}$ when we quantize the theory.

Quantum mechanically, we use the Hamiltonian (12.21) to discuss quantum motion, using the commutators

$$[x_i, p_j] = i\hbar\delta_{ij}\,, \quad [x_i, x_j] = [p_i, p_j] = 0\,. \tag{12.22}$$

Expanding Eq. (12.21), we have

$$H = \frac{\boldsymbol{p}^2}{2m} - \frac{e}{2mc}(\boldsymbol{p}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{p}) + \frac{e^2}{2mc^2}\boldsymbol{A}^2 + e\phi\,. \tag{12.23}$$

Note that the ordering of $\boldsymbol{p}$ and $\boldsymbol{A}$ matters here, as they may not commute.

12.2.3 Gauge Invariance in Quantum Mechanics

The Schrödinger equation is

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= \left[\frac{\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2}{2m} + e\phi\right]\psi \\ &= \left[\frac{\left(-i\hbar\nabla - \frac{e}{c}\boldsymbol{A}\right)^2}{2m} + e\phi\right]\psi\,. \end{aligned} \tag{12.24}$$

How should we change the wavefunction so that the theory remains invariant under gauge transformations?

Let

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\Lambda\,, \quad \phi' = \phi - \frac{1}{c}\frac{\partial\Lambda}{\partial t}\,, \tag{12.25}$$

which is a gauge transformation, where $\Lambda(\boldsymbol{x}, t)$ is a scalar function. Then we have

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= \left[\frac{\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2}{2m} + e\phi\right]\psi \\ &= \left[\frac{\left(-i\hbar\nabla - \frac{e}{c}\boldsymbol{A}' + \frac{e}{c}\nabla\Lambda\right)^2}{2m} + e\phi' + \frac{e}{c}\frac{\partial\Lambda}{\partial t}\right]\psi\,. \end{aligned} \tag{12.26}$$

We can rearrange this equation to reach

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi' - \frac{e}{c}\frac{\partial\Lambda}{\partial t}\right)\psi = \frac{\left(-i\hbar\nabla - \frac{e}{c}\boldsymbol{A}' + \frac{e}{c}\nabla\Lambda\right)^2}{2m}\psi\,. \tag{12.27}$$

We can compensate the changes from the gauge transformation by letting

$$\psi = \psi' e^{-ie\Lambda/\hbar c}\,, \tag{12.28}$$

which leaves us with

$$\left(i\hbar\frac{\partial}{\partial t} - e\phi'\right)\psi' = \frac{\left(-i\hbar\nabla - \frac{e}{c}\boldsymbol{A}'\right)^2}{2m}\psi'\,, \tag{12.29}$$

matching the form of the Schrödinger equation in Eq. (12.24).

Thus, the full gauge transformation is

$$\begin{aligned} \boldsymbol{A}' &= \boldsymbol{A} + \nabla\Lambda\,, \\ \phi' &= \phi - \frac{1}{c}\frac{\partial\Lambda}{\partial t}\,, \\ \psi &= \psi' e^{-ie\Lambda/\hbar c}\,. \end{aligned} \tag{12.30}$$

All observables must remain invariant under this transformation.

12.2.4 Aharonov–Bohm Effect

Consider a hollow cylindrical shell with a solenoid through the middle, such that there is a flux

$$\Phi = \int_{\text{core}} \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{S} \tag{12.31}$$

through the core, but nowhere in the shell. Suppose that some charged particles are constrained to move within the shell, where there is no magnetic field. Classically, the magnetic field does not affect the motion of the charged particles, because they experience no forces. What about in quantum mechanics?

In order to formulate the quantum motion, we need to determine the vector potential $\boldsymbol{A}$ (note that the scalar potential is zero in this case). Even though the magnetic field is zero in the shell where the particles move, the vector potential cannot be chosen to vanish in this region. To see this, consider taking the line integral of $\boldsymbol{A}$ around a closed contour C lying within the shell and enclosing the core. Using Stokes' theorem, we see that this line integral is given by

$$\oint_C \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{\ell} = \int_\Sigma (\nabla\times\boldsymbol{A})\cdot \mathrm{d}\boldsymbol{S} = \int_\Sigma \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{S} = \Phi\,, \tag{12.32}$$

where Σ is the area bounded by the contour C, i.e., $C = \partial\Sigma$. Because there is nonzero flux through the core, this integral cannot be zero. Thus, we cannot set $\boldsymbol{A}$ to zero everywhere in the shell.

Lecture 13 (Oct. 25, 2017)

13.1 Aharonov–Bohm Effect

Recall from last lecture the setup for the Aharonov–Bohm effect: we have a cylindrical shell through which runs a solenoid, producing a flux

$$\Phi = \int_{\text{core}} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} \tag{13.1}$$

through the core of the shell. There is no magnetic field inside the cylindrical shell, only through its core. As we saw, even though the magnetic field within the shell vanishes, it is impossible to choose the vector potential $\boldsymbol{A}$ to vanish inside the cylindrical shell. To see why, we took a closed contour C within the shell that enclosed the core, and noted that

$$\oint_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\ell} = \int_\Sigma (\nabla \times \boldsymbol{A}) \cdot \mathrm{d}\boldsymbol{S} = \int_\Sigma \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = \Phi\,, \tag{13.2}$$

where Σ is the area bounded by the contour C, i.e., $C = \partial\Sigma$. This shows us that $\boldsymbol{A}$ cannot vanish everywhere within the shell.

In particular, let us consider the contour C to be a circular loop of radius R. The line integral above has contributions only from the azimuthal component of the vector potential. Thus, we should be able to produce the magnetic field coming from the solenoid by taking a vector potential of the form

$$\boldsymbol{A} = A_\theta \hat{\boldsymbol{e}}_\theta\,. \tag{13.3}$$

By cylindrical symmetry, it is reasonable to take A_θ to be independent of θ. We then have

$$\oint_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\ell} = A_\theta(2\pi R) = \Phi\,, \tag{13.4}$$

which yields

$$A_\theta = \frac{\Phi}{2\pi R}\,. \tag{13.5}$$

Thus, a vector potential that produces the desired magnetic field is

$$\boldsymbol{A}(r) = \frac{\phi}{2\pi r}\hat{\boldsymbol{e}}_\theta\,, \tag{13.6}$$

where r is the radial cylindrical coordinate. Note that there are infinitely many vector potentials that will produce the same magnetic field, all related to one another by gauge transformations.

Because we have a nonzero vector potential, there will be nontrivial quantum mechanical behavior. Thinking in terms of the propagator

$$K(\boldsymbol{x}_f, t_f; \boldsymbol{x}_i, t_i) = \int [\mathcal{D}\boldsymbol{x}]\; e^{iS[\boldsymbol{x}(t)]/\hbar}\,. \tag{13.7}$$

The action is

$$\begin{aligned} S[\boldsymbol{x}(t)] &= \int \mathrm{d}t\; L(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) \\ &= \int \mathrm{d}t\; \left(\frac{1}{2}m\dot{\boldsymbol{x}}^2 + \frac{e}{c}\boldsymbol{A}\cdot\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}\right) \\ &= \int \mathrm{d}t\; \frac{1}{2}m\dot{\boldsymbol{x}}^2 + \frac{e}{c}\int \mathrm{d}\boldsymbol{x}\cdot\boldsymbol{A}(\boldsymbol{x}) \\ &= S_0 + \frac{e}{c}\int \mathrm{d}\boldsymbol{x}\cdot\boldsymbol{A}(\boldsymbol{x})\,, \end{aligned} \tag{13.8}$$

where S_0 is the action without the presence of any magnetic field. Thus, the amplitude to propagate from $(\boldsymbol{x}_i, t_i)$ to $(\boldsymbol{x}_f, t_f)$ is equal to the amplitude with no magnetic field multiplied by the phase

$$\exp\left(\frac{ie}{\hbar c}\int_{\boldsymbol{x}_i}^{\boldsymbol{x}_f} \mathrm{d}\boldsymbol{x}\cdot\boldsymbol{A}\right). \tag{13.9}$$

So far this statement is general; let us now specialize to the case of the Aharonov–Bohm setup. Consider two points A and B distributed symmetrically around the cylindrical shell, and both lying in the same plane orthogonal to the axis of the cylindrical shell. We consider two types paths from A to B within this plane: those passing around the left of the cylindrical shell, and those passing around the right of the cylindrical shell. Note that our choice of a particular path around the left of the shell does not affect the phase accumulated along this path. To see this, note that we could choose two such paths and combine them to form a closed loop which encloses no magnetic flux; the closed line integral of $\boldsymbol{A}$ along this loop therefore vanishes, showing that the phase picked up on the two paths must be equal. In other words, for any two paths P_L and P'_L passing around the left, we have

$$\int_{P_\mathrm{L}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x} = \int_{P'_\mathrm{L}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x}\,. \tag{13.10}$$

This is similarly true for the paths around the right of the shell.

The amplitudes when traveling along the left and right paths P_L and P_R are therefore

$$\begin{aligned} a(P_\mathrm{L}) &= a_0 e^{\frac{ie}{\hbar c}\int_{P_\mathrm{L}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x}}\,,\\ a(P_\mathrm{R}) &= a_0 e^{\frac{ie}{\hbar c}\int_{P_\mathrm{R}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x}}\,. \end{aligned} \tag{13.11}$$

If we consider the difference of the additional phase factors, we have

$$\int_{P_\mathrm{R}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x} - \int_{P_\mathrm{L}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x} = \oint \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x} = \Phi \neq 0\,. \tag{13.12}$$

The total propagator is thus

$$\begin{aligned} K &= K_\mathrm{L} e^{\frac{ie}{\hbar c}\int_{P_\mathrm{L}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x}} + K_\mathrm{R} e^{\frac{ie}{\hbar c}\int_{P_\mathrm{R}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x}}\\ &= e^{\frac{ie}{\hbar c}\int_{P_\mathrm{L}} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{x}}\left(K_\mathrm{L} + K_\mathrm{R} e^{ie\Phi/\hbar c}\right). \end{aligned} \tag{13.13}$$

The probability of propagating from point A to point B is then

$$\mathrm{Prob}(A\to B) = |K_\mathrm{L}|^2 + |K_\mathrm{R}|^2 + 2K_\mathrm{L}K_\mathrm{R}\cos\left(\frac{e\Phi}{\hbar c}\right). \tag{13.14}$$

This probability oscillates as a function of the flux Φ, even though classical trajectories do not see the magnetic field. We see that this oscillation depends only on the flux Φ, and not on our choice of gauge, as is required by gauge invariance.

We have derived this effect using the path integral approach. We could have also reached this result by using the Schödinger equation. The energy levels and wavefunctions for a particle in a ring threaded by flux Φ are dependent on Φ.

This result leads us to the following question: the vector potential is a (redundant) mathematical construct that we use to describe the physics, but the actually physical quantities are the electromagnetic fields. How can it be the case that the vector potential makes an imprint on the physics, when we know that the vector potential by itself is not gauge invariant? To reconcile the

Aharonov–Bohm effect and the non-reality of the vector potential, we must realize that the physics of a quantum particle is sensitive not only to the local magnetic field, but also the global properties of the magnetic field. This is a departure from the classical intuition: quantum mechanics requires nonlocal effects. Formulating the theory in the way that we have, using the vector potential, allows us to define the theory only locally and find the same results. We have formulated these nonlocal effects in terms of a purely local description, but the price we pay is that there is no unique description of the formalism, because we have gauge redundancy.

The same approach can be seen in quantum field theory. With quantum electrodynamics, we describe the physics of charged particles and photons. We all know that two charged particles interact via the Coulomb interaction, which is nonlocal; we can think of the introduction of photons as taming this nonlocality in the same way as the introduction of the vector potential tamed the nonlocality in the Aharonov–Bohm setup.

13.2 Magnetic Monopoles

Suppose that magnetic monopoles exist. This requires a modification to Gauss's law for magnetism,

$$\nabla \cdot \boldsymbol{B} = 0 \to \nabla \cdot \boldsymbol{B} = 4\pi\rho_M \,, \tag{13.15}$$

where ρ_M is the density of magnetic charge (as well as modifications to Faraday's law and the Lorentz force law). A magnetic monopole is a single magnetic charge of strength g. If there is a magnetic monopole of magnetic charge g at the origin, then

$$\boldsymbol{B} = \frac{g}{r^2}\hat{\boldsymbol{e}}_r \,. \tag{13.16}$$

What is the fate of a quantum mechanical particle of electric charge e moving in the presence of a magnetic monopole? We immediately realize a problem. We typically describe the magnetic field in quantum mechanics via the vector potential, but the existence of the vector potential is predicated on the fact that the divergence of the magnetic field vanishes. In a world where magnetic monopoles exist, the magnetic field does not have a representation as the curl of a vector potential.

Let's think about this problem in the presence of a single magnetic monopole sitting at the origin. In this situation, $\nabla \cdot \boldsymbol{B} = 0$ everywhere except at the origin. The implication is that if we exclude the point at the origin, then we can *locally* find a vector potential $\boldsymbol{A}$ such that $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. To see more clearly why we cannot describe the magnetic field everywhere with a vector potential, consider a sphere of radius r centered at the origin. Due to the presence of the magnetic monopole at the origin, we see that

$$\int_{\text{sphere}} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 4\pi g \,. \tag{13.17}$$

If we were able to write $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ with $\boldsymbol{A}$ well-defined *everywhere*, then we would have

$$\int_{\text{sphere}} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = \int \nabla \cdot \boldsymbol{B}\, \mathrm{d}V = \int \nabla \cdot (\nabla \times \boldsymbol{A})\, \mathrm{d}V = 0 \,, \tag{13.18}$$

which yields a contradiction.

Consider a circle C on the surface of the sphere at angle θ. The circle divides the sphere into two caps; the magnetic flux through the upper cap of the sphere bounded by this circle is

$$\Phi = 2\pi \int \frac{g}{r^2} r^2\, \mathrm{d}(\cos\theta) = 2\pi g(1 - \cos\theta) \,. \tag{13.19}$$

If there exists a vector potential

$$\boldsymbol{A} = A_\phi \hat{\boldsymbol{e}}_\phi \,, \tag{13.20}$$

then

$$\oint_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\ell} = 2\pi r \sin\theta A_\phi \,. \tag{13.21}$$

On the other hand,

$$\oint_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\ell} = \int_\Sigma \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 2\pi g(1 - \cos\theta) \,, \tag{13.22}$$

where Σ is the upper cap bounded by C, i.e., $C = \partial\Sigma$. Thus, we have

$$A_\phi = \frac{g(1 - \cos\theta)}{r \sin\theta} \,. \tag{13.23}$$

Note that this ratio is well-behaved as $\theta \to 0$, because the numerator goes to zero more quickly than the denominator, but the ratio blows up as $\theta \to \pi$.

Lecture 14 (Oct. 30, 2017)

14.1 Magnetic Monopoles

Last time, we considered a magnetic field with a magnetic monopole configuration, and began to approach describing the quantum mechanics of a charged particle moving in such a magnetic field. The presence of the monopole violates the condition $\nabla \boldsymbol{B} = 0$, which means that we cannot globally define a vector potential $\boldsymbol{A}$ such that $\boldsymbol{B} = \nabla \times \boldsymbol{A}$.

We began by considering an isolated magnetic monopole of strength g located at the origin, which gives a magnetic field

$$\boldsymbol{B} = \frac{g}{r^2}\hat{\boldsymbol{e}}_r \,. \tag{14.1}$$

Consider a sphere of radius r, and a closed circular contour C at θ on the sphere. The magnetic flux through the upper cap Σ_+ bounded by C is

$$\Phi_C = 2\pi g(1 - \cos\theta) \,. \tag{14.2}$$

If we choose

$$\boldsymbol{A} = A_\phi \hat{\boldsymbol{e}}_\phi \,, \tag{14.3}$$

then we have

$$\oint_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{\ell} = 2\pi r \sin\theta A_\phi = \int_{\Sigma_+} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 2\pi g(1 - \cos\theta) \,, \tag{14.4}$$

which gives us

$$A_\phi = \frac{g(1 - \cos\theta)}{r \sin\theta} \,. \tag{14.5}$$

Where is this potential well-defined? We see that it diverges at $r = 0$, but we expected this, because the magnetic field is divergent at $r = 0$. As $\theta \to 0$, the cosine in the numerator approaches 1 more quickly than the sine in the denominator approaches 0, so A_ϕ is well-defined at $\theta = 0$. However, it blows up at $\theta = \pi$. Let us rename this vector potential, as

$$A_\phi^+ = \frac{g(1 - \cos\theta)}{r \sin\theta} \,. \tag{14.6}$$

In calculating this vector potential, we made the arbitrary choice to consider the magnetic flux through the upper cap bounded by C instead of the lower cap. If we instead consider the flux through the lower cap, we would find a vector potential $\boldsymbol{A} = A_\phi^- \hat{\boldsymbol{e}}_\phi$, where

$$A_\phi^- = -\frac{g(1 + \cos\theta)}{r \sin\theta} \,. \tag{14.7}$$

This vector potential is not well-defined at $\theta = 0$, but is well-defined at $\theta = \pi$.

Thus, we cannot use A_ϕ^+ at the south pole, and we cannot use A_ϕ^- at the north pole, but everywhere else we could equally well choose either expression for the vector potential. Because both of these vector potentials describe the same electromagnetic fields in these regions, they must differ by the gradient of some function. Indeed, we see that

$$\boldsymbol{A}^+ - \boldsymbol{A}^- = \left(A_\phi^+ - A_\phi^-\right)\hat{\boldsymbol{e}}_\phi = \frac{2g}{r \sin\theta}\hat{\boldsymbol{e}}_\phi = 2g\nabla\phi \,. \tag{14.8}$$

Thus, as expected, these two vector potentials differ from one another by a gauge transformation.

We have found that, in order to describe the magnetic monopole field in terms of vector potentials, we can write $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, where

$$\boldsymbol{A} = \begin{cases} \boldsymbol{A}^+ = \frac{g(1-\cos\theta)}{r\sin\theta}\hat{\boldsymbol{e}}_\phi\,, & \text{for } 0 \le \theta < \frac{\pi}{2} + \epsilon\,, \\ \boldsymbol{A}^- = -\frac{g(1+\cos\theta)}{r\sin\theta}\hat{\boldsymbol{e}}_\phi\,, & \text{for } \frac{\pi}{2} - \epsilon < \theta \le \pi\,, \end{cases} . \tag{14.9}$$

In the region where $\boldsymbol{A}^+$ and $\boldsymbol{A}^-$ equal $\boldsymbol{A}$, they are well-defined. In the overlap region, the differ by a gauge transformation.

14.1.1 QM of a Charged Particle Moving in a Magnetic Monopole Field

Consider a particle of electric charge e. For $0 \le \theta < \frac{\pi}{2} + \epsilon$, let the wavefunction be $\psi^+(r, \theta, \phi)$, and for $\frac{\pi}{2} - \epsilon < \theta \le \pi$, let the wavefunction be $\psi^-(r, \theta, \phi)$. Assume that we have already determined these by solving the Schrödinger equation. In the overlap region, ψ^+ and ψ^- must be related by a gauge transformation. We saw in a previous lecture that if we make a gauge transformation

$$\boldsymbol{A}^+ = \boldsymbol{A}^- + 2g\nabla\phi\,, \tag{14.10}$$

then the wavefunction must change by

$$\psi^+ = \psi^- e^{2ieg\phi/\hbar c}\,. \tag{14.11}$$

Now, we note that the wavefunctions ψ^+ and ψ^- must be single-valued when $\phi \to \phi + 2\pi$. This requires that

$$\frac{2eg}{\hbar c} = n \in \mathbb{Z}\,, \tag{14.12}$$

i.e.,

$$g = \frac{nhc}{2\pi e}\,, \quad n \in \mathbb{Z}\,. \tag{14.13}$$

This is the *Dirac quantization condition*: the magnetic monopole field strength must be quantized, with the quantum a function of the electric charge.

We can gain intuition about this result with a vague classical analogue. Note, however, that this quantization is a purely quantum effect, and so cannot be fully explained classically. Let's consider the charged particle moving in the field of an magnetic monopole from a classical point of view. Consider an electric monopole of strength e and a magnetic monopole of strength g, both static, and displaced from one another by distance d along the z-axis. We can then ask about the total angular momentum stored in the electromagnetic field; this information is contained in the Poynting vector, which is proportional to $\boldsymbol{E} \times \boldsymbol{B}$. Purely by symmetry, we conclude that the angular momentum must be directed along the z-axis. If we carry out the calculation, we find that the total angular momentum is independent of the distance d, and is proportional to eg. In quantum mechanics, we know that angular momentum is quantized. If we require the total angular momentum we found to be an integer multiple of $\frac{\hbar}{2}$, then we recover the Dirac quantization condition. Incidentally, this gives us a model of what is "spinning" in a spin-$\frac{1}{2}$ particle: a bound state of a bosonic magnetic monopole and a bosonic electric monopole has spin-$\frac{1}{2}$. The naïve statistics of this bound state would be bosonic, because both the electric and magnetic monopoles are bosonic; however, the interactions between the two charges lead to a change in the statistics.

14.2 Charged Particle in a Uniform Magnetic Field

Consider a particle of electric charge e moving in a uniform magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$ in three dimensions. The Hamiltonian is

$$H = \frac{\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2}{2m}\,. \tag{14.14}$$

Only $B_z \neq 0$, and so we can always choose $A_z = 0$ and A_x, A_y independent of z. The Hamiltonian then becomes

$$H = \frac{p_z^2}{2m} + \frac{\Pi_x^2 + \Pi_y^2}{2m}\,, \tag{14.15}$$

where

$$\Pi_{x,y} = p_{x,y} - \frac{e}{c}A_{x,y} \tag{14.16}$$

are the kinematic momenta in the x- and y-directions. Note that $[p_z, H] = 0$, so we can label the eigenstates by p_z. We can then write the Hamiltonian in the form

$$H = \frac{p_z^2}{2m} + H_{2\mathrm{d}}\,, \tag{14.17}$$

with

$$H_{2\mathrm{d}} = \frac{\Pi_x^2 + \Pi_y^2}{2m}\,, \tag{14.18}$$

and we only have to determine the spectrum of $H_{2\mathrm{d}}$.

The trick is to notice that Π_x and Π_y have a simple commutation relation,

$$[\Pi_x, \Pi_y] = \left[-i\hbar\partial_x - \frac{e}{c}A_x, -i\hbar\partial_y - \frac{e}{c}A_y\right] = \frac{eB}{c}i\hbar\,. \tag{14.19}$$

Thus, these two kinematic momenta (appropriately rescaled) are canonically conjugate variables, and the Hamiltonian $H_{2\mathrm{d}}$ looks like the sum of squares of canonically conjugate variables, which is the Hamiltonian of the simple harmonic oscillator in one dimension. More precisely, let

$$X = \frac{c\Pi_x}{eB}\,, \quad P = \Pi_y\,, \tag{14.20}$$

Then, $[X, P] = i\hbar$, and

$$H_{2\mathrm{d}} = \frac{P^2}{2m} + \frac{1}{2m}\left(\frac{eB}{c}\right)^2 X^2 = \frac{P^2}{2m} + \frac{1}{2}m\omega_c^2 X^2\,, \tag{14.21}$$

with

$$\omega_c = \frac{eB}{mc}\,. \tag{14.22}$$

This is the one-dimensional SHO Hamiltonian, with frequency ω_c, known as the *cyclotron frequency*.

As an aside, the classical motion of a charged particle a uniform magnetic field is described by circular orbits in a plane orthogonal to the magnetic field. Matching the centrifugal force with the force from the magnetic field, we have

$$\frac{mv^2}{R} = \frac{evB}{c}\,, \tag{14.23}$$

which gives a radius of

$$R = \frac{mvc}{eB}\,, \tag{14.24}$$

known as the *cyclotron radius*. The time period of the orbit is

$$T = \frac{2\pi R}{v} = \frac{2\pi}{\omega_c}\,, \tag{14.25}$$

with ω_c the cyclotron frequency. This is the classical origin of the cyclotron frequency.

Now, we have the Hamiltonian

$$H_{\text{2d}} = \frac{P^2}{2m} + \frac{1}{2}m\omega_c^2 X^2 \tag{14.26}$$

with $[X, P] = i\hbar$. We can immediately conclude that the energy levels are

$$E_n^{(\text{2d})} = \hbar\omega\left(n + \frac{1}{2}\right). \tag{14.27}$$

The three-dimensional energy levels of the full Hamiltonian H are then

$$E_n^{(\text{3d})}(p_z) = \frac{p_z^2}{2m} + \hbar\omega\left(n + \frac{1}{2}\right). \tag{14.28}$$

However, we are not done, because we do not know the degeneracies of these energy levels. We will find that the spectrum is highly degenerate.

14.2.1 Degeneracy

Why is there degeneracy in the spectrum? One way to understand the degeneracy is to notice that we can define new coordinates in the problem,

$$R_x := x + \frac{c}{eB}\Pi_y\,, \quad R_y := y - \frac{c}{eB}\Pi_x\,. \tag{14.29}$$

Note that $[R_x, R_y] = -\frac{c}{eB}i\hbar$. Thus, R_x and R_y are canonically conjugate up to a multiplicative factor. Furthermore, we note that

$$\begin{aligned}[R_x, \Pi_x] &= [x, \Pi_x] + \frac{c}{eB}[\Pi_y, \Pi_x] \\ &= i\hbar - \frac{c}{eB}\left(\frac{eB}{c}i\hbar\right) \\ &= 0\,.\end{aligned} \tag{14.30}$$

Similarly, we find that $[R_x, \Pi_y] = 0$, and more generally,

$$[R_i, \Pi_j] = 0\,. \tag{14.31}$$

Thus,

$$[R_i, H] = 0\,. \tag{14.32}$$

We have two operators that each commute with the Hamiltonian, but they do not commute with one another. Recall from a previous homework that when this is the case, the Hamiltonian must be degenerate.

What is the physical meaning of these coordinates R_i? Recall that classically, the particle undergoes circular motion in the presence of the uniform magnetic field. Classically, the vector $\boldsymbol{R} = (R_x, R_y)$ is the center of the cyclotron orbit: if (x, y) are the time-dependent coordinates

of the particle moving in its circular orbit, and we take $\Pi_i = mv_i$, then (R_x, R_y) are the time-independent coordinates of the center of the orbit. This point is called the *guiding center*. We see that in quantum mechanics, the coordinates of the guiding center do not commute with one another in the presence of the magnetic field. The size of the cyclotron orbit will be fixed such that the magnetic flux through the orbit yields one of the quantized energies, but the location of the orbit is not fixed, which leads to the degeneracy. These degenerate energy levels are called *Landau levels*.

Lecture 15 (Nov. 1, 2017)

15.1 Charged Particle in a Uniform Magnetic Field

Last time, we discussed the quantum mechanics of a charged particle moving in a uniform magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$. We decomposed the Hamiltonian as

$$H = \frac{p_z^2}{2m} + H_{\text{2d}}\,, \tag{15.1}$$

with

$$H_{\text{2d}} = \frac{\Pi_x^2 + \Pi_y^2}{2m}\,. \tag{15.2}$$

The Hamiltonian H_{2d} has spectrum

$$E_n^{(\text{2d})} = \hbar\omega_c\left(n + \frac{1}{2}\right), \tag{15.3}$$

with the cyclotron frequency given by

$$\omega_c = \frac{eB}{mc}\,. \tag{15.4}$$

We then wanted to determine the degeneracy of these energy levels, called *Landau levels*. We defined the *guiding center coordinates*,

$$R_x := x + \frac{c}{eB}\Pi_y\,, \quad R_y := y - \frac{c}{eB}\Pi_x\,. \tag{15.5}$$

These variables are conjugate, up to a multiplicative factor,

$$[R_x, R_y] = -\frac{c}{eB}i\hbar = -i\ell_B^2\,, \tag{15.6}$$

and both commute with the Hamiltonian. As we showed on the homework, this implies that the Hamiltonian is degenerate. Here, ℓ_B is the *magnetic length*, which is the length scale such that if we cut out a circular disk of radius ℓ_B orthogonal to the magnetic field, then the flux through the disk is on the order of one flux quantum:

$$\pi B\ell_B^2 = \frac{hc}{2e} \sim \frac{hc}{e} = \Phi_0\,, \tag{15.7}$$

where Φ_0 is the flux quantum.

The guiding center coordinates are the coordinates of the center of the cyclotron orbit, which are a constant of the motion. This statement is equivalent to the statement that the guiding center coordinates commute with the Hamiltonian. In the quantum theory, the size of the orbit is set by the only length scale in the problem, the magnetic length. We expect that the degeneracy is the number of flux quanta passing through the sample, $\sim \frac{BA}{\Phi}$, where A is the area of the sample.

Now we will calculate the degeneracy explicitly. We pick a particular gauge to work in, for convenience. We choose

$$A_y = Bx\,, \quad A_x = 0\,. \tag{15.8}$$

This is known as *Landau gauge*. In this gauge,

$$\begin{aligned} H_{\text{2d}} &= \frac{p_x^2}{2m} + \frac{\left(p_y - \frac{eBx}{c}\right)^2}{2m} \\ &= \frac{p_x^2}{2m} + \frac{1}{2}m\omega_c^2\left(x - \frac{cp_y}{eB}\right)^2. \end{aligned} \tag{15.9}$$

In this gauge, $[p_y, H] = 0$, so we can label states by their p_y eigenvalue. For any fixed p_y, we get a one-dimensional simple harmonic oscillator of frequency ω_c, which reproduces the spectrum

$$E_n^{(2\mathrm{d})} = \hbar\omega_c\left(n + \frac{1}{2}\right). \tag{15.10}$$

The energy eigenfunctions take the form

$$\psi_{n,p_y} = e^{ip_y y/\hbar}\phi_n\left(x - \frac{cp_y}{eB}\right). \tag{15.11}$$

This is a plane wave multiplied by a shifted SHO wavefunction. It is now clear that states with different values of p_y will be degenerate. The degeneracy question then becomes the question of how many distinct values of p_y exist for a system of a given area.

Consider a sample of lengths L_x and L_y along the x- and y-directions, respectively, with periodic boundary conditions along the y-direction. We then must have

$$e^{ip_y L_y/\hbar} = 1\,, \tag{15.12}$$

which gives us quantized momentum modes along the y-direction,

$$p_y = \frac{2\pi m\hbar}{L_y}\,, m \in \mathbb{Z}\,. \tag{15.13}$$

Each of the degenerate wavefunctions $\psi_{n,p_y}(x,y)$ has a modulus $\left|\phi_n\left(x - \frac{cp_y}{eB}\right)\right|$ and is centered at

$$x = \left(\frac{c}{eB}\right)\left(\frac{2\pi m\hbar}{L_y}\right). \tag{15.14}$$

However, we cannot arbitrarily shift these SHO groundstate wavefunctions, because the sample has a finite length L_x in the x-direction. Thus, the y-direction momentum is quantized because of the finite length in the y-direction, and has a finite number of possible values because of the finite length in the x-direction. The number of independent states is then

$$g = \frac{L_x}{\left(\frac{c}{eB}\right)\left(\frac{2\pi\hbar}{L_y}\right)} = \frac{eBL_xL_y}{2\pi\hbar c} = \frac{eBA}{hc} = \frac{BA}{\Phi_0}\,, \tag{15.15}$$

as we expected.

In classical mechanics, we can prove that there is no diamagnetism. In classical statistical mechanics, everything is determined by the partition function,

$$Z \propto \int \prod \frac{\mathrm{d}x_i\,\mathrm{d}p_i}{(2\pi\hbar)^3}\, e^{-\beta\left(\sum_i \frac{\boldsymbol{p}_i^2}{2m} + V(\{\boldsymbol{x}_i\})\right)}. \tag{15.16}$$

In the presence of a magnetic field, the Hamiltonian changes,

$$Z[\boldsymbol{A}] \propto \int \prod \frac{\mathrm{d}x_i\,\mathrm{d}p_i}{(2\pi\hbar)^3}\, e^{-\beta\left(\sum_i \frac{\left(\boldsymbol{p}_i - \frac{e}{c}\boldsymbol{A}(\boldsymbol{x}_i)\right)^2}{2m} + V(\{\boldsymbol{x}_i\})\right)}. \tag{15.17}$$

However, we can simply shift the momenta as $\boldsymbol{p}_i' = \boldsymbol{p}_i - \frac{e}{c}\boldsymbol{A}(\boldsymbol{x}_i)$, which then gives us $Z[A] = Z[A = 0]$. Thus, classically, the magnetic field does not affect the statistical mechanics. However, we have just seen that in quantum mechanics, there is an effect. This effect is purely quantum mechanical.

We might wonder how a two-dimensional problem of motion in the x, y-plane turned into a one-dimensional SHO. The electron is localized around the guiding center coordinates. However, we see that the two coordinates do not commute with one another. Thus, there is only one real coordinate in the problem, and the other coordinate is its conjugate variable. This is why we are only solving a one-dimensional problem at the end of the day. Note that because R_x and R_y do not commute, we cannot be in a simultaneous eigenstate of both, so the location of guiding center itself must be smeared out.

15.2 Composite Systems

We now consider composite systems, which are systems composed of two (or more) subsystems. Suppose we have two systems A and B, with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. We want to obtain the Hilbert space of the composite space $A + B$, which is called the *tensor product* of $\mathcal{H}_A$ and $\mathcal{H}_B$.

Consider a basis $|\{\phi_i\}\rangle$ of $\mathcal{H}_A$ and a basis $|\{\chi_j\}\rangle$ of $\mathcal{H}_B$. We will define new vectors $|\phi_i\rangle \otimes |\chi_j\rangle$ that live in a new Hilbert space, denoted $\mathcal{H}_{A+B} = \mathcal{H}_A \otimes \mathcal{H}_B$. These $|\phi_i\rangle \otimes |\chi_j\rangle$ will provide a basis for the Hilbert space $\mathcal{H}_{A+B}$, meaning that any state in $\mathcal{H}_{A+B}$ can be decomposed as

$$|\psi_{A+B}\rangle = \sum_{i,j} \psi_{ij} |\phi_i\rangle \otimes |\chi_j\rangle \,. \tag{15.18}$$

Often times, we will omit the tensor product symbol $\otimes$ for convenience, when the meaning is clear. Thus, we will write $|\phi_i\rangle|\chi_j\rangle = |\phi_i\rangle \otimes |\chi_j\rangle$.

We can then write down operators on $\mathcal{H}_{A+B}$ and formulate the quantum theory as we have before. There are classes of operators that act on only one of the two subsystems, such as $\mathcal{O} = \mathcal{O}_A \otimes \mathbb{1}_B$, but in general we can have operators that act nontrivially on both subsystems at once.

Consider, in particular, a system of two spin-$\frac{1}{2}$ particles A and B. Some possible states in the composite Hilbert space can be written as a single tensor product, such as

$$|\uparrow_A\rangle \otimes |\uparrow_B\rangle \,, \quad |\downarrow_A\rangle \otimes |\uparrow_B\rangle \,, \quad |\uparrow_A\rangle \otimes \left[\frac{1}{\sqrt{2}}(|\uparrow_B\rangle + |\downarrow_B\rangle)\right] . \tag{15.19}$$

There are clearly an infinite number of such states. There are also more interesting states (entangled states) that cannot be written as a single tensor product, such as

$$\frac{1}{\sqrt{2}}(|\uparrow_A\rangle \otimes |\downarrow_B\rangle - |\downarrow_A\rangle \otimes |\uparrow_B\rangle) \,. \tag{15.20}$$

Some of the operators on this Hilbert space act only on one of the subsystems, such as $\sigma_A^z \otimes \mathbb{1}_B := \sigma_A^z$ (the second way of writing this is a notational convenience). However, there are more general operators that act nontrivially on both subsystems simultaneously, such as $\sigma_A^x \otimes \sigma_B^y := \sigma_A^x \sigma_B^y$.

15.2.1 Quantum Entanglement

Two subsystems A and B of a quantum mechanical system can be entangled with one another. For example, consider again the system of two spin-$\frac{1}{2}$ particles A and B. Some states, such as

$$|\uparrow_A\rangle \otimes |\downarrow_B\rangle \,, \quad |\uparrow_A\rangle \otimes |\downarrow_B\rangle \,, \tag{15.21}$$

have the property that if we restrict our view to only one of the subsystems, then the system is in a well-defined state of that subsystem. By contrast, if we consider a state such as

$$\frac{1}{\sqrt{2}}(|\uparrow_A\rangle \otimes |\downarrow_B\rangle \pm |\downarrow_A\rangle \otimes |\uparrow_B\rangle) \,, \tag{15.22}$$

then we see that neither spin by itself is in a definite quantum state, despite the fact that the composite system is in a definite state. This is an *entangled state*.

Entanglement is a statement about the relation of subsystems of a system to one another. It does not make sense to ask if the state of a system "is entangled." Instead, we can make statements about whether two subsystems are entangled with one another. If we have subsystems A and B of some composite system, then an unentangled state is precisely one that can be factored as a tensor product of states of the two subsystems,

$$|\psi_{A+B}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \,. \tag{15.23}$$

An entangled state is one that cannot be factorized in this way.

Lecture 18 (Nov. 13, 2017)

18.1 Symmetries in Quantum Mechanics

A symmetry is a physical operation we can perform on the system that leaves the physics unchanged. As an example, consider a free particle,

$$H_{\text{free}} = \frac{p^2}{2m}\,. \tag{18.1}$$

This Hamiltonian does not depend on position, so we can translate $x \to x + a$. Under this transformation, H_{free} is unchanged and $[x + a, p] = i\hbar$, implying that both H and the commutation algebra are unchanged. Thus, translation is a symmetry of the free-particle system. This Hamiltonian has several other symmetries. *Inversion* (or *parity*), given by $x \to -x, p \to -p$, preserves $[x, p] = i\hbar$ and H_{free}, and so is a symmetry. *Time reversal*, which sends $t \to -t$, is also a symmetry; we will discuss its realization in quantum mechanics later. The system is also invariant under Galilean transformations, even though Galilean boosts $p \to p + mv_0$ seem to change the Hamiltonian. Galilean boosts change the phase in the path integral by something that does not depend on the path, so the probabilities do not change, but the amplitudes do; this example is more subtle, and so we will stick to discussions of the other symmetries of the system.

18.1.1 Symmetry Transformations

We define a *symmetry operation* as a linear transformation $U\colon |a\rangle \to |a'\rangle$ such that all results of measurement are preserved. This means that

$$\left|\langle b'|a'\rangle\right|^2 = |\langle b|a\rangle|^2 \tag{18.2}$$

for all $a, b \in \mathcal{H}$. From the definition of the primed kets, this gives

$$\left|\langle b|U^\dagger U|a\rangle\right|^2 = |\langle b|a\rangle|^2\,. \tag{18.3}$$

Wigner's theorem tells us that if U satisfies this condition, then U must be unitary or anti-unitary; an *anti-unitary* operator $\tilde{U}$ is one that satisfies

$$\tilde{U}|a\rangle = KU|U\rangle = (U|a\rangle)^\dagger\,, \tag{18.4}$$

where K is the complex conjugation operator, and U is some unitary operator. If we assume that U is linear or *anti-linear*, meaning that

$$U\left(\sum_a c_a|a\rangle\right) = \sum_a c_a^* U|a\rangle\,, \tag{18.5}$$

then we can show Wigner's theorem fairly simply (this will appear on your homework). Wigner's theorem can be proved without making this assumption, but the proof is more subtle.

18.1.2 Continuous Symmetries and Conservation Laws

Consider a symmetry operation described by a unitary operator U. This is a symmetry of the Hamiltonian if H is unchanged by the action of U, i.e.

$$H = U^\dagger H U\,, \tag{18.6}$$

i.e., $[H, U] = 0$. If the symmetry transformation can be continuously built up as a series of infinitesimal transformations starting from the identity operator $\mathbb{1}$, then we call it a *continuous symmetry*. Translation and rotation are examples of continuous symmetries: any translation or rotation can be built up from the identity by doing a sequence of infinitesimal translations or rotations. Symmetries that cannot be built up in this way are called *discrete symmetries*. Parity is an example of a discrete symmetry (in even dimensions, we can construct parity as a rotation, but it is nonetheless possible that the Hamiltonian is symmetric under parity but not rotations, so parity is still discrete).

Let us now consider continuous symmetries. All of the information about a continuous symmetry is contained in its infinitesimal form, because we can always build up a continuous transformation from its infinitesimal version. As we have seen previously, an infinitesimal unitary operator can be expressed in terms of a Hermitian operator: if we expand an infinitesimal unitary operator U as

$$U = 1 - \frac{i\epsilon}{\hbar} G + O\big(\epsilon^2\big)\,, \tag{18.7}$$

then the statement $U^\dagger U = \mathbb{1}$ implies $G^\dagger = G$, i.e., G is Hermitian. The statement that U is a symmetry is $U^\dagger H U = H$, i.e.,

$$\left(1 + \frac{i\epsilon}{\hbar} G + O\big(\epsilon^2\big)\right) H \left(1 - \frac{i\epsilon}{\hbar} G + O\big(\epsilon^2\big)\right) = H\,, \tag{18.8}$$

which gives us $[G, H] = 0$. Because G is Hermitian, it corresponds to some observable, and we now have

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \frac{1}{i\hbar}[G, H] = 0\,, \tag{18.9}$$

so we see that G is conserved. This leads us to the following general statement:

Theorem 4 (Noether's Theorem)**.** *For every continuous symmetry of the Hamiltonian in quantum mechanics, there is a corresponding conserved quantity. Conversely, if some observable G is conserved, then $[G, H] = 0$, and we can define unitary operators*

$$U(\theta) = e^{i\theta G/\hbar}\,, \tag{18.10}$$

which will satisfy $U(\theta)^\dagger H U(\theta) = H$, showing that $U(\theta)$ is a continuous symmetry of the Hamiltonian.

18.1.3 Translations

As an example, consider again translation, $x \to x + a$. The translation operator is

$$T(a) = e^{-ipa/\hbar}\,. \tag{18.11}$$

This satisfies $(T(a))^\dagger = (T(a))^{-1}$, i.e., the translation operator is unitary. It has the following properties:

- $(T(a))^{-1} = T(-a)$,
- $T(a')T(a'') = T(a' + a'')$,
- $T^\dagger(a) x T(a) = x + a$.

For infinitesimal a, we can expand the translation operator as

$$T(a) \approx 1 - \frac{ia}{\hbar} p \,, \tag{18.12}$$

where p is the Hermitian *generator* of $T(a)$. If a system has translation symmetry, then momentum is conserved, from Noether's theorem. In d dimensions, we generalize to translations $x_i \to x_i + a_i$, with $i = 1, \ldots, d$. Then, the unitary translation operator is

$$T(\{a_i\}) = \prod_i T_i(a_i) = \prod_i e^{-ip_i a_i/\hbar} \,. \tag{18.13}$$

It is a geometric fact that translations in different directions commute, $[T_i(a_i), T_j(a_j)] = 0$, for all i, j. This tells us that the Hermitian generators commute, $[p_i, p_j] = 0$, for all i, j. We can define momentum as the observable that corresponds to the Hermitian generator of translations; then taking a to be infinitesimal in the identity

$$e^{ipa/\hbar} x e^{-ipa/\hbar} = x + a \tag{18.14}$$

gives us the commutation relation $[x, p] = i\hbar$.

18.1.4 Time Translations

For a closed system, time translation is a symmetry, and the corresponding unitary operator is the time-evolution operator,

$$U(t) = e^{-iHt/\hbar} \,. \tag{18.15}$$

The Hamiltonian H is the generator of this symmetry, and so the fact that the system is invariant under time translation implies the conservation of energy,

$$\frac{\mathrm{d}H}{\mathrm{d}t} = 0 \,. \tag{18.16}$$

18.1.5 Rotations

If we have a d-dimensional system with position coordinate x_i, then a rotation is given by

$$x_i \to x_i' = \sum_{j=1}^{d} R_{ij} x_j = R_{ij} x_j \,, \tag{18.17}$$

such that the scalar product is preserved,

$$\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}' \cdot \boldsymbol{y}' \,. \tag{18.18}$$

Here we are using the Einstein summation convention that repeated indices are summed over. The preservation of the scalar product can be written as

$$R_{ij} R_{ik} x_j y_k = x_i y_i \,, \tag{18.19}$$

which implies

$$R_{ij} R_{ik} = \delta_{jk} \,, \tag{18.20}$$

i.e., $RR^{\mathrm{T}} = \mathbb{1}$. That is, R is an orthogonal matrix.

The orthogonality condition, $RR^{\mathrm{T}} = \mathbb{1}$, implies that $(\det R)^2 = 1$, meaning that $\det R = \pm 1$. We see then that there are two classes of rotations. Matrices with $\det R = -1$ are not continuously connected to the identity; they involve inversions $x_i \to -x_i$ combined with an ordinary (orientation-preserving) rotation. (Note that in even dimensions, parity is orientation-preserving.) For now, we will stick to considering rotations R that have $\det R = +1$. Each rotation R is represented in quantum mechanics by a unitary operator $\mathcal{D}(R)$ that acts on the system's Hilbert space.

Lecture 19 (Nov. 15, 2017)

19.1 Rotations

Recall that rotations are transformations of the form $x_i \to R_{ij}x_j$ (using Einstein summation notation), where R is an orthogonal matrix, $R^{\mathrm{T}}R = \mathbb{1}$. This R is called a *rotation matrix*. For now, we will restrict to rotations R with $\det R = +1$ (*orientation-preserving* or *proper* rotations).

Every rotation R of space corresponds to a unitary operator $\mathcal{D}(R)$ on the Hilbert space, which satisfies

$$\mathcal{D}(R_1)\mathcal{D}(R_2) = \mathcal{D}(R_1R_2)\,. \tag{19.1}$$

We will discuss this composition property more later. A quantum state transforms as $|\alpha\rangle \to |\alpha_R\rangle$ under this rotation, such that

$$|\alpha_R\rangle = \mathcal{D}(R)|\alpha\rangle\,. \tag{19.2}$$

For a vector operator V_i, with $i = 1, \ldots, d$, we require

$$\langle\beta_R|V_i|\alpha_R\rangle = R_{ij}\langle\beta|V_j|\alpha\rangle\,. \tag{19.3}$$

Expanding the transformed bra and ket, this is

$$\langle\beta|\mathcal{D}^\dagger(R)V_i\mathcal{D}(R)|\alpha\rangle = R_{ij}\langle\beta|V_j|\alpha\rangle\,. \tag{19.4}$$

This is true for any states $|\alpha\rangle, |\beta\rangle$, which implies that the operators must be equal:

$$\mathcal{D}^\dagger(R)V_i\mathcal{D}(R) = R_{ij}V_j \tag{19.5}$$

holds as an operator equation.

Consider the infinitesimal rotation $R = 1 - \omega$. The orthogonality condition, $R^{\mathrm{T}}R = \mathbb{1}$, then implies that $\omega^{\mathrm{T}} = -\omega$. Thus, ω is a real, antisymmetric matrix. We can then expand $\mathcal{D}(R)$ in the form

$$\mathcal{D}(R) \approx 1 - \frac{i}{2\hbar}\sum_{ij}\omega_{ij}J_{ij} + O\big(\omega^2\big)\,. \tag{19.6}$$

This expansion identifies the objects $J_{ij} = -J_{ji}$ as the Hermitian generators of rotations. Note that the antisymmetry of J_{ij} follows from the antisymmetry of ω_{ij}.

Let us now specialize to three dimensions, $d = 3$. In this case, we can write

$$\mathcal{D}(R) = 1 - \frac{i}{\hbar}(J_{12}\omega_{12} + J_{23}\omega_{23} + J_{31}\omega_{31}) + O\big(\omega^2\big)\,, \tag{19.7}$$

where we have used the antisymmetry of J_{ij} to group terms. We define

$$J_1 := J_{23}\,, \quad J_2 := J_{31}\,, \quad J_3 := J_{12}\,, \tag{19.8}$$

i.e.,

$$J_i = \frac{1}{2}\epsilon_{ijk}J_{jk}\,, \tag{19.9}$$

where ϵ_{ijk} is the totally antisymmetric symbol with $\epsilon_{123} = +1$, known as the *Levi–Civita symbol*. We can similarly define

$$\theta_1 := \omega_{23}\,, \quad \theta_2 := \omega_{31}\,, \quad \theta_3 := \omega_{12}\,, \tag{19.10}$$

i.e.,

$$\theta_i = \frac{1}{2}\epsilon_{ijk}\omega_{jk}\,, \quad \omega_{ij} = \epsilon_{ijk}\theta_k\,. \tag{19.11}$$

Then, we have

$$\mathcal{D}(R) = 1 - \frac{i}{\hbar}\theta_k J_k + O\big(\theta^2\big)\,. \tag{19.12}$$

Note that

$$R_{ij} = \delta_{ij} - \epsilon_{ijk}\theta_k + O\big(\theta^2\big)\,. \tag{19.13}$$

Thus,

$$x_i \to x_i' = R_{ij}x_j = (\delta_{ij} - \epsilon_{ijk}\theta_k)x_j = x_i - \epsilon_{ijk}x_j\theta_k\,, \tag{19.14}$$

i.e.,

$$\boldsymbol{x} \to \boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{\theta}\times\boldsymbol{x}\,. \tag{19.15}$$

Thus, the meaning of $\boldsymbol{\theta}$ is that $\boldsymbol{x}$ is rotated by an angle $|\boldsymbol{\theta}|$ about the $\boldsymbol{\theta}$-direction.

We now *define* J_k to be the components of the angular momentum. First, we will derive the commutation relations of J_k with any vector operator V_i. We start with the equation

$$\mathcal{D}^\dagger(R)V_i\mathcal{D}(R) = R_{ij}V_j \tag{19.16}$$

and take $R = 1 - \omega$ with ω infinitesimal. The left-hand side then becomes

$$\left(1 + \frac{i\theta_k J_k}{\hbar}\right)V_i\left(1 - \frac{i\theta_\ell J_\ell}{\hbar}\right) = V_i + \frac{i\theta_k}{\hbar}[J_k, V_i]\,, \tag{19.17}$$

while the right-hand side becomes

$$R_{ij}V_j = V_i - \epsilon_{ijk}V_j\theta_k\,. \tag{19.18}$$

Thus, we conclude that

$$[J_k, V_i] = i\hbar\epsilon_{kij}V_j\,. \tag{19.19}$$

We can then use a combination of rotations to deduce the angular momentum algebra, via

$$\mathcal{D}(R_1)\mathcal{D}(R_2) = \mathcal{D}(R_1R_2)\,. \tag{19.20}$$

In particular, this composition rule implies that

$$\mathcal{D}(R_{\boldsymbol{\phi}})\mathcal{D}(R_{\boldsymbol{\theta}})\mathcal{D}\Big(R_{\boldsymbol{\phi}}^{-1}\Big) = \mathcal{D}\Big(R_{\boldsymbol{\phi}}R_{\boldsymbol{\theta}}R_{\boldsymbol{\phi}}^{-1}\Big)\,. \tag{19.21}$$

The rotation $R_{\boldsymbol{\phi}}R_{\boldsymbol{\theta}}R_{\boldsymbol{\phi}}^{-1}$ can be written as a single rotation $R_{\boldsymbol{\theta}'}$ for some $\boldsymbol{\theta}'$. As $\boldsymbol{\theta}$ itself is a vector, for $\boldsymbol{\phi}$ infinitesimal, we have

$$\boldsymbol{\theta}' = \boldsymbol{\theta} + \boldsymbol{\phi}\times\boldsymbol{\theta}\,. \tag{19.22}$$

If we take $\boldsymbol{\theta}$ to be infinitesimal, we have

$$\mathcal{D}(R_{\boldsymbol{\theta}}) = 1 - \frac{i\theta_k J_k}{\hbar} + O\big(\theta^2\big)\,, \tag{19.23}$$

so (19.21) becomes

$$\theta_k\mathcal{D}(R_{\boldsymbol{\phi}})J_k\mathcal{D}\Big(R_{\boldsymbol{\phi}}^{-1}\Big) = \theta_k' J_k \tag{19.24}$$

The left-hand side of this equation, for infinitesimal $\boldsymbol{\phi}$ is

$$\theta_k\left(1 - \frac{i\phi_j J_j}{\hbar}\right)J_k\left(1 + \frac{i\phi_\ell J_\ell}{\hbar}\right) = \theta_k J_k - \frac{i\theta_k\phi_j}{\hbar}[J_j, J_k] + \cdots\,, \tag{19.25}$$

while the right-hand side is

$$\theta_k J_k + \epsilon_{jk\ell}\phi_j\theta_k J_\ell + \cdots , \tag{19.26}$$

which leads us to conclude that

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k . \tag{19.27}$$

This is the angular momentum commutation algebra. Note that this matches the commutation relation for a vector operator with the angular momentum operator, so this shows that the angular momentum operator is a vector.

In general, we can write the angular momentum as

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S} , \tag{19.28}$$

with $\boldsymbol{L} = \boldsymbol{x}\times\boldsymbol{p}$ is the orbital angular momentum and $\boldsymbol{S}$ is an internal property that commutes with $\boldsymbol{x}$, $\boldsymbol{p}$, etc. We can check that $\boldsymbol{L}$ on its own satisfies the angular momentum commutation algebra, so the operator $\boldsymbol{J}$ will satisfy the angular momentum algebra if $\boldsymbol{S}$ does. The operator $\boldsymbol{S}$ is the spin operator.

If the Hamiltonian is rotationally invariant, then $[J_i, H] = 0$, which implies that

$$\frac{\mathrm{d}J_i}{\mathrm{d}t} = 0 , \tag{19.29}$$

and so angular momentum is conserved.

19.1.1 Eigensystem of Angular Momentum

Let us now understand the implications of the commutation algebra

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k . \tag{19.30}$$

You will show on the homework that

$$\left[\boldsymbol{J}^2, J_i\right] = 0 . \tag{19.31}$$

This means that we can diagonalize $\boldsymbol{J}^2$ and one component of the angular momentum, say J_z, simultaneously. We can then label the eigenstates of J_z by $|j, m\rangle$, with

$$\boldsymbol{J}^2|j,m\rangle = a|j,m\rangle , \quad J_z|j,m\rangle = b|j,m\rangle , \tag{19.32}$$

for some eigenvalues a, b. The meanings of the values j and m will become apparent shortly.

It is useful to define the *ladder operators*

$$J_\pm = J_x \pm iJ_y , \tag{19.33}$$

which satisfy

$$[J_+, J_-] = 2\hbar J_z , \quad [J_z, J_\pm] = \pm\hbar J_\pm , \quad \left[\boldsymbol{J}^2, J_\pm\right] = 0 . \tag{19.34}$$

Using these commutation relations, we see that

$$\begin{aligned} J_z(J_\pm|j,m\rangle) &= (J_\pm J_z \pm \hbar J_\pm)|j,m\rangle \\ &= (b \pm \hbar)(J_\pm|j,m\rangle) . \end{aligned} \tag{19.35}$$

Thus, $J_\pm|j,m\rangle$ is also an eigenstates of J_z with eigenvalue $b \pm \hbar$.

We can write

$$\begin{aligned}\boldsymbol{J}^2 &= J_x^2 + J_y^2 + J_z^2 \\ &= J_z^2 + \frac{1}{2}(J_+J_- + J_-J_+) \\ &= J_z^2 + \frac{1}{2}\Big(J_+J_+^\dagger + J_-J_-^\dagger\Big)\,,\end{aligned} \tag{19.36}$$

which tells us that $\boldsymbol{J}^2 - J_z^2$ is positive semi-definite,

$$\big\langle j,m\big|\boldsymbol{J}^2 - J_z^2\big|j,m\big\rangle \geq 0\,. \tag{19.37}$$

This implies that $a - b^2 \geq 0$ for all eigenstates. For a fixed a, this means that $|b|$ has a maximum value $b_{\max}$. This seems to be in conflict with the statement that we can use $J_\pm$ to raise or lower the eigenvalue arbitrarily. We conclude that at $b = +b_{\max}$ the state must be annihilated by J_+, and similarly, at $b = -b_{\max}$ the state must be annihilated by J_-.

Call $|\max\rangle$ the state with $b = +b_{\max}$. Then, we have

$$J_+|\max\rangle = 0\,, \tag{19.38}$$

which implies

$$J_-J_+|\max\rangle = 0\,. \tag{19.39}$$

Expanding the ladder operators, this becomes

$$(J_x - iJ_y)(J_x + iJ_y)|\max\rangle = \big(\boldsymbol{J}^2 - J_z^2 - \hbar J_z\big)|\max\rangle = 0\,. \tag{19.40}$$

This gives us

$$a - b_{\max}^2 - \hbar b_{\max} = 0\,, \tag{19.41}$$

i.e.,

$$a = b_{\max}(b_{\max} + \hbar)\,. \tag{19.42}$$

Repeating this argument for the state $|\min\rangle$ with $b = -b_{\max}$ yields the same result.

We now note that, because J_+ increases the eigenvalue b, and this eigenvalue is bounded above by $b_{\max}$, we must be able to reach $|\max\rangle$ from $|\min\rangle$ by repeatedly applying J_+. Say we can reach $|\max\rangle$ from $|\min\rangle$ by n applications of J_+. This implies that

$$b_{\max} = -b_{\max} + n\hbar\,, \tag{19.43}$$

so

$$b_{\max} = \frac{n\hbar}{2} = j\hbar\,, \tag{19.44}$$

with $j \in \frac{1}{2}\mathbb{Z}$. We can then read off the eigenvalues for the state $|j,m\rangle$,

$$a = \hbar^2 j(j+1)\,, \quad b = m\hbar\,, \tag{19.45}$$

and see that m can take any of the $2j+1$ values $-j, -j+1, \ldots, j-1, j$.

Lecture 20 (Nov. 20, 2017)

20.1 Matrix Elements of Angular Momentum Operators

Assume that we have normalized $|j,m\rangle$. By definition, since these are eigenstates of $\boldsymbol{J}^2$ and J_z, these operators are diagonal, with

$$\begin{aligned}\langle j',m'|\boldsymbol{J}^2|j,m\rangle &= j(j+1)\hbar^2\delta_{jj'}\delta_{mm'}\,,\\ \langle j',m'|J_z|j,m\rangle &= m\hbar\delta_{jj'}\delta_{mm'}\,.\end{aligned} \tag{20.1}$$

Now we only need to compute the matrix elements of J_x and J_y. We will make use of the identity

$$\langle j,m|J_+J_-|j,m\rangle = \langle j,m|\boldsymbol{J}^2 - J_z^2 - \hbar J_z|jm\rangle = \hbar^2\big(j(j+1)-m^2-m\big)\,. \tag{20.2}$$

We also know that

$$J_+|j,m\rangle = c^{(+)}_{j,m}|j,m+1\rangle\,. \tag{20.3}$$

We still need to determine the coefficient $c^{(+)}_{j,m}$. Using (20.2) and $J_- = J_+^\dagger$, we see that

$$\left|c^{(+)}_{j,m}\right|^2 = \hbar^2(j(j+1)-m(m+1)) = \hbar^2(j-m)(j+m+1)\,. \tag{20.4}$$

We will choose (by convention) for $c^{(+)}_{j,m}$ to be real and positive, which then gives us

$$J_+|j,m\rangle = \hbar\sqrt{(j-m)(j+m+1)}|j,m+1\rangle\,. \tag{20.5}$$

A similar argument gives us

$$J_-|j,m\rangle = \hbar\sqrt{(j+m)(j-m+1)}|j,m-1\rangle\,. \tag{20.6}$$

Packaging these results together, we have

$$\langle j',m'|J_\pm|j,m\rangle = \hbar\sqrt{(j\mp m)(j\pm m+1)}\delta_{jj'}\delta_{m',m\pm1}\,. \tag{20.7}$$

We see that when $m=+j$, the state is annihilated by J_+, and when $m=-j$, the state is annihilated by J_-, exactly as required.

We can now easily calculate the matrix elements of J_x and J_y, as they are linear combinations of $J_\pm$. This allows us to explicitly write matrix representations of $\boldsymbol{J}$ for fixed j. As an example, consider $j=1$. In this case,

$$J_z = \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}, \tag{20.8}$$

where the basis is ordered as $m=+1,0,-1$. Using the matrix elements for $J_\pm$, we can write down

$$\begin{aligned}J_x &= \frac{J_+ + J_-}{2} = \frac{\hbar}{2}\begin{pmatrix}0&\sqrt{2}&0\\\sqrt{2}&0&\sqrt{2}\\0&\sqrt{2}&0\end{pmatrix},\\ J_y &= \frac{J_x - J_y}{2i} = \frac{\hbar}{2}\begin{pmatrix}0&-\sqrt{2}i&0\\\sqrt{2}i&0&-\sqrt{2}i\\0&\sqrt{2}i&0\end{pmatrix}.\end{aligned} \tag{20.9}$$

For any fixed j, we can now also construct matrix elements of the finite rotation operator $\mathcal{D}(R)$. For a fixed j, the angular momentum operators are $(2j+1)\times(2j+1)$ Hermitian matrices, and so the rotation matrix $\mathcal{D}(R)$ will be $(2j+1)\times(2j+1)$ unitary matrices. Consider a rotation by angle θ about the axis $\hat{\boldsymbol{n}}$. First, we construct an infinitesimal rotation by $\delta\theta$ about the axis $\hat{\boldsymbol{n}}$:

$$\mathcal{D}(R_{\delta\theta}) = 1 - \frac{i}{\hbar}(\boldsymbol{J}\cdot\hat{\boldsymbol{n}})\delta\theta + O\big(\delta\theta^2\big)\,. \tag{20.10}$$

We can then build up a finite rotation by multiplying infinitesimal rotations N times and taking $N\to\infty$,

$$\begin{aligned}\mathcal{D}(R(\theta,\hat{\boldsymbol{n}})) &= \lim_{N\to\infty}\big(\mathcal{D}(\theta/N,\hat{\boldsymbol{n}})\big)^N \\ &= \lim_{N\to\infty}\left(1 - \frac{i}{\hbar}(\boldsymbol{J}\cdot\hat{\boldsymbol{n}})\frac{\theta}{N} + O\big((\theta/N)^2\big)\right)^N \\ &= e^{-i\theta\boldsymbol{J}\cdot\hat{\boldsymbol{n}}/\hbar}\,.\end{aligned} \tag{20.11}$$

20.2 Rotation Groups

In general, symmetry operations form a mathematical structure called a *group*.

Definition 1. A set $\{g_i\} := G$ forms a group if there exists a multiplication operation $\cdot$ such that:

- for all $g_1, g_2 \in G$, $g_1\cdot g_2 \in G$;
- multiplication is associative: for all $g_1, g_2, g_3 \in G$,

$$(g_1\cdot g_2)\cdot g_3 = g_1\cdot(g_2\cdot g_3)\,; \tag{20.12}$$

- there exists $1\in G$ such that

$$1\cdot g = g\cdot 1 = g\,; \tag{20.13}$$

- for every $g\in G$, there exists $g^{-1}\in G$ such that

$$g\cdot g^{-1} = g^{-1}\cdot g = 1\,. \tag{20.14}$$

For example, the set of 3×3 orthogonal matrices forms a group under the operation of matrix multiplication. This group is known as the *orthogonal group* O(3). If we restrict our attention to rotations with $\det R = +1$, then we get a subgroup (a subset that is itself also a group under matrix multiplication) called the *special orthogonal group* SO(3).

There is another pertinent group to the discussion of rotations, which is the group of 2×2 unitary matrices satisfying $\det U = 1$. Any matrix of this kind can be written in the form

$$U(a,b) = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{20.15}$$

with $|a|^2 + |b|^2 = 1$. Such matrices form a group under matrix multiplication, known as the *special unitary group* SU(2).

In quantum mechanics, rotations are represented by unitary operators $\mathcal{D}(R)$. These unitary operators must themselves form a group, because the product of two of these operators must give another, the composition is associative, there is an identity operation, and $\mathcal{D}\big(R^{-1}\big)$ is the inverse of $\mathcal{D}(R)$. In particular, we can see that the $\mathcal{D}(R)$ must satisfy the same group identities as the group of rotations R.

The matrix elements of $\mathcal{D}(R)$ are

$$\mathcal{D}^{(j)}_{m',m}(R) = \langle j, m'|e^{-i\theta \boldsymbol{J}\cdot\hat{\boldsymbol{n}}/\hbar}|j,m\rangle\,. \tag{20.16}$$

Note that we have chosen the same j for both the bra and ket; this is because $\mathcal{D}(R)$ commutes with $\boldsymbol{J}^2$, so a rotation cannot change the value of $\boldsymbol{J}^2$. As a result, in the full ket space, the (infinite-dimensional) matrix for $\mathcal{D}(R)$ is block diagonal in the (j,m) basis, where each block corresponds to one value of j. The block corresponding to j will be a $(2j+1)\times(2j+1)$ submatrix.

An explicit set of matrices that satisfies the same group identities as a group is called a *(linear) group representation*. Thus, the $\mathcal{D}(R)$ form a representation of SO(3). In particular, this representation is called a *completely reducible* representation, because it can be written in block diagonal form. Each block on the diagonal of $\mathcal{D}(R)$, taken as a standalone matrix $\mathcal{D}^{(j)}_{m',m}$ for fixed j, is an *irreducible* representation (sometimes referred to as an *irrep*), as it has no invariant subspaces.

Consider $j = 1/2$. In this case, m and m' take on the values $\pm\frac{1}{2}$. The rotation operator matrix elements are

$$\mathcal{D}^{(1/2)}_{m',m}(\theta,\hat{\boldsymbol{n}}) = \langle m'|e^{-i\theta\boldsymbol{\sigma}\cdot\hat{\boldsymbol{n}}/2}|m\rangle\,. \tag{20.17}$$

That is, the rotation operators are

$$\begin{aligned}\mathcal{D}^{(1/2)}(\theta,\hat{\boldsymbol{n}}) &= e^{-i\theta\boldsymbol{\sigma}\cdot\hat{\boldsymbol{n}}/2}\\ &= \cos\frac{\theta}{2} - i(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{n}})\sin\frac{\theta}{2}\,.\end{aligned} \tag{20.18}$$

Explicitly,

$$\mathcal{D}^{(1/2)}(\theta,\hat{\boldsymbol{n}}) = \begin{pmatrix} \cos\frac{\theta}{2} - in_z\sin\frac{\theta}{2} & (-in_x - n_y)\sin\frac{\theta}{2} \\ (-in_x + n_y)\sin\frac{\theta}{2} & \cos\frac{\theta}{2} + in_z\sin\frac{\theta}{2}\end{pmatrix}. \tag{20.19}$$

This may seem unusual, because this is an SU(2) matrix. The operators $\mathcal{D}^{(1/2)}(\theta,\hat{\boldsymbol{n}})$ form a representation of SU(2), even though we originally set out to find a representation of SO(3).

Rotations of 3D space can be characterized either by an SO(3) matrix R (a three-dimensional irrep) or by an SU(2) matrix U (a two-dimensional irrep). But SO(3) and SU(2) are not in one-to-one correspondence. Suppose we consider a rotation by $\theta = 2\pi$ along some axis $\hat{\boldsymbol{n}}$. For SO(3),

$$R(2\pi,\hat{\boldsymbol{n}}) = \mathbb{1}_3\,, \tag{20.20}$$

while for SU(2),

$$\mathcal{D}^{(1/2)}(2\pi,\hat{\boldsymbol{n}}) = e^{-i\pi\boldsymbol{\sigma}\cdot\hat{\boldsymbol{n}}} = -\mathbb{1}_2\,. \tag{20.21}$$

The result is that for any spin-$\frac{1}{2}$ system, the phase of the wavefunction is rotated by π under a physical 2π rotation.

Lecture 21 (Nov. 22, 2017)

21.1 SO(3) versus SU(2)

As we saw, if we study only the representations of SO(3), we will only find representations with integer spin. However, by considering the Lie algebra $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, we found representations with half-integer spin. These should properly be considered as representations of SU(2), which is a double cover of SO(3), but has the same Lie algebra. Linear representations of SU(2) include both integer and half-integer spins, while linear representations of SO(3) include only integer spins. As representations of SO(3), the half-integer spin representations are what are called *projective* representations, meaning that they satisfy the group structure only up to a phase:

$$\mathcal{D}(R_1)\mathcal{D}(R_2) = e^{i\xi(R_1,R_2)}\mathcal{D}(R_1R_2)\,. \tag{21.1}$$

The elements of SU(2) are in two-to-one correspondence with the elements of SU(3), which is why we say that SU(2) is a double cover of SO(3). The two SU(2) matrices U and $-U$ both correspond to the same SO(3) matrix R. This can be seen directly from the expression

$$R_{ij} = \frac{1}{2}\,\mathrm{tr}\Big(\sigma_j U^\dagger \sigma_i U\Big)\,, \tag{21.2}$$

which expresses a 3×3 rotation matrix R in terms of a 2×2 unitary matrix U. On the homework, you will prove that R defined in this way is a proper rotation matrix.

21.2 Addition of Angular Momentum

Given two systems 1 and 2, with angular momenta j_1 and j_2, respectively, what are the possible total angular momenta for the combined system? We first must choose a basis for the combined Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. We will choose the basis

$$|j_1, m_1; j_2, m_2\rangle = |j_1.m_1\rangle \otimes |j_2, m_2\rangle\,. \tag{21.3}$$

We know how to write down operators on this tensor product Hilbert space. For example, the total angular momentum is

$$\boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2 := \boldsymbol{J}_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \boldsymbol{J}_2\,, \tag{21.4}$$

where the rightmost expression makes the meaning completely clear, and the middle expression is a convenient shorthand that suppresses the products with the identity. Note that this combined operator satisfies the angular momentum commutation algebra, $[J_i, J_j] = i\epsilon_{ijk}J_k$. (If we took the difference of two angular momenta, rather than the sum, then the result would *not* satisfy the angular momentum commutation algebra.) Note also that

$$\Big[J_1^i, J_2^j\Big] = 0\,, \tag{21.5}$$

because the spin operators on one subsystem act trivially on the other subsystem.

We have

$$\boldsymbol{J}^2 = \boldsymbol{J}_1^2 + \boldsymbol{J}_2^2 + 2\boldsymbol{J}_1 \cdot \boldsymbol{J}_2\,, \tag{21.6}$$

from which we see that

$$\big[\boldsymbol{J}^2, \boldsymbol{J}_{a,z}\big] \neq 0\,, \tag{21.7}$$

where $a = 1, 2$. Thus, we see that the eigenvalue of $\boldsymbol{J}^2$ is not a good quantum number in the basis we have chosen, in the sense that the basis we have chosen is not an eigenbasis of $\boldsymbol{J}^2$.

However, because $\boldsymbol{J}$ satisfies the angular momentum commutation algebra, we know that $\boldsymbol{J}^2$ and J_z can be simultaneously diagonalized. What are the allowed eigenvalues of $\left(\boldsymbol{J}^2, J_z\right)$, given J_1 and J_2? As an example, if we consider two spin-$\frac{1}{2}$ systems, there are four states in the combined Hilbert space, which can be written in the form

$$|j_1 = \frac{1}{2}, m_1, j_2 = \frac{1}{2}, m_2\rangle = |{+}{+}\rangle, |{+}{-}\rangle, |{-}{+}\rangle, |{-}{-}\rangle\,. \tag{21.8}$$

Clearly,

$$J_z|{+}{+}\rangle = |{+}{+}\rangle\,, \tag{21.9}$$

which implies that $m = +1$ for $|{+}{+}\rangle$. Similarly, we find $m = -1$ for $|{-}{-}\rangle$, and $m = 0$ for both $|{+}{-}\rangle$ and $|{-}{+}\rangle$. These organize into a singlet $j = 0$,

$$\frac{1}{\sqrt{2}}(|{+}{-}\rangle - |{-}{+}\rangle)\,, \tag{21.10}$$

and a triplet $j = 1$,

$$\begin{gathered} |{+}{+}\rangle\,, \\ \frac{1}{\sqrt{2}}(|{+}{-}\rangle + |{-}{+}\rangle)\,, \\ |{-}{-}\rangle\,. \end{gathered} \tag{21.11}$$

In general, we want to know what the coefficients are for the change of basis from $|j_1, m_1; j_2, m_2\rangle$ to the new basis $|j, m; j_1, j_2\rangle$, i.e., the coefficients

$$\langle j_1, m_1; j_2, m_2 | j, m; j_1, j_2\rangle\,. \tag{21.12}$$

These are known as *Clebsch–Gordon coefficients*. We now discuss some general properties of Clebsch–Gordon coefficients:

1. The Clebsch–Gordon coefficients are zero unless $m = m_1 + m_2$. To see this, note that

   $$J_z - J_{1,z} - J_{2,z} = 0 \tag{21.13}$$

   as an operator. In particular,

   $$(J_z - J_{1,z} - J_{2,z})|j, m; j_1, j_2\rangle = 0\,. \tag{21.14}$$

   Then,

   $$\langle j_1, m_1; j_2, m_2|(J_z - J_{1,z} - J_{2,z})|j, m; j_1, j_2\rangle = 0\,. \tag{21.15}$$

   Acting with J_z to the right and with $J_{i,z}$ to the left, this gives us

   $$(m - m_1 - m_2)\langle j_1, m_1; j_2, m_2|j, m; j_1, j_2\rangle = 0\,. \tag{21.16}$$

   Thus, the Clebsch–Gordon coefficient must vanish unless $m = m_1 + m_2$.

2. The Clebsch–Gordon coefficients vanish unless $|j_1 - j_2| \leq j \leq j_1 + j_2$. As a consistency check, we can see that this condition gives the correct number of states. In the basis $|j_1, m_1; j_2.m_2\rangle$, we can see that the total number of states is

   $$(2j_1 + 1)(2j_2 + 1)\,, \tag{21.17}$$

because subsystem i has $2j_i + 1$ possible states. We can similarly count the number of states in the basis $|j, m; j_1, j_2\rangle$, assuming that $|j_1 - j_2| \leq j \leq j_1 + j_2$. For each j, there are $2j + 1$ states, and j runs between $|j_1 - j_2|$ and $j_1 + j_2$. We further assume that every j in this range occurs, but with two successive j values differing by 1 and not by $\frac{1}{2}$. This gives the total number of states as

$$\sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) = (2j_1+1)(2j_2+1)\,, \tag{21.18}$$

where we have assumed $j_1 \geq j_2$, without loss of generality.

3. If j_1 and j_2 are both half-integer (by which we mean an integer plus $\frac{1}{2}$) or both integer, then j is integer. To see this, note that under a 2π rotation, both subsystems will acquire the same phase (-1 if they are both half-integer, or $+1$ if they are both integer), and so the total system acquires no phase. Thus, the combined system has integer angular momentum.

   Using the same type of argument, we see that the converse is also true: if only one of (j_1, j_2) is half-integer, then j is half-integer.

21.3 Discrete Symmetries

We now move on to discuss parity (a.k.a. inversion). Let Π be the operator that implements the parity transformation. By this, we mean that

$$|\alpha\rangle \to \Pi|\alpha\rangle := |\alpha_\Pi\rangle \tag{21.19}$$

under the parity transformation $\boldsymbol{x} \to -\boldsymbol{x}$. We require that

$$\langle\alpha_\Pi|\boldsymbol{x}|\alpha_\Pi\rangle = -\langle\alpha|\boldsymbol{x}|\beta\rangle\,, \tag{21.20}$$

which implies that

$$\Pi^\dagger \boldsymbol{x} \Pi = -\boldsymbol{x}\,. \tag{21.21}$$

In other words,

$$\{\boldsymbol{x}, \Pi\} = 0\,, \tag{21.22}$$

where $\{\cdot,\cdot\}$ is the anticommutator.

Two successive inversions do nothing, i.e., $\Pi^2 = \mathbb{1}$. This implies that the eigenvalues of Π are ± 1. States with eigenvalue $+1$ are called *even* under parity, and states with eigenvalue -1 are called *odd*. Operators that are odd under parity must anticommute with Π. For example, $\boldsymbol{p}$ is odd, as $\boldsymbol{p} \to -\boldsymbol{p}$ under parity, so

$$\{\boldsymbol{p}, \Pi\} = 0\,. \tag{21.23}$$

What about angular momentum? For a general rotation (not necessarily a proper rotation),

$$\Pi R(\theta, \hat{\boldsymbol{n}}) = R(\theta, \hat{\boldsymbol{n}})\Pi\,. \tag{21.24}$$

That is, parity and rotation commute with one another. This implies that $[\Pi, \boldsymbol{J}] = 0$. In general, $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$. We can see that $\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p}$ is even under parity, because $\boldsymbol{x}$ and $\boldsymbol{p}$ are each odd. Because $\boldsymbol{J}$ and $\boldsymbol{L}$ are both even under parity, we conclude that $\boldsymbol{S}$ must also be even under parity.

Lecture 22 (Nov. 27, 2017)

22.1 Parity

22.1.1 Some Standard Terminology

When we refer to a *scalar*, we mean an observable that is invariant under rotations and even under parity. Examples include $\boldsymbol{x}^2$ and $\boldsymbol{p}^2$. There is a different type of object, called a *pseudoscalar*, that is invariant under rotations, but odd under parity. An example of a pseudoscalar is the product $\boldsymbol{S}\cdot\boldsymbol{x}$; this is invariant under rotations, but odd under parity because $\boldsymbol{S}$ is parity even and $\boldsymbol{x}$ is parity odd.

A *vector* is an object that transforms as a vector under rotations and is odd under parity. Examples are $\boldsymbol{x}$ and $\boldsymbol{p}$. A *pseudovector* is an object that transforms as a vector under rotations, but is even under parity. Examples of pseudovectors are $\boldsymbol{L}$, $\boldsymbol{S}$, and $\boldsymbol{J}$.

22.1.2 Wavefunctions Under Parity

Eigenstates of parity satisfy

$$\Pi|\psi\rangle = \pm|\psi\rangle\,, \tag{22.1}$$

as we know that Π has eigenvalues ±1 only. If we take the matrix element with a position ket, then we find

$$\langle\boldsymbol{x}|\Pi|\psi\rangle = \pm\langle\boldsymbol{x}|\psi\rangle = \pm\psi(\boldsymbol{x})\,. \tag{22.2}$$

On the other hand, we can have the parity operator act on the position ket, giving

$$\langle\boldsymbol{x}|\Pi|\psi\rangle = \langle-\boldsymbol{x}|\psi\rangle = \psi(-\boldsymbol{x})\,. \tag{22.3}$$

Thus, wavefunctions of parity eigenstates satisfy

$$\psi(-\boldsymbol{x}) = \pm\psi(\boldsymbol{x})\,. \tag{22.4}$$

We refer to such wavefunctions as *even* $(+)$ or *odd* $(-)$.

22.1.3 Momentum and Angular Momentum

As we have seen, $[\boldsymbol{p},\Pi]\neq0$. Thus, we cannot simultaneously diagonalize the momentum and parity operators, i.e., momentum eigenstates are not, in general, parity eigenstates.

As an example, consider the free particle

$$H = \frac{p^2}{2m}\,. \tag{22.5}$$

The energy eigenstates

$$\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} \tag{22.6}$$

are not parity eigenstates. However, $[\Pi,H]=0$, which means that we can choose energy eigenstates that are also parity eigenstates in this case. Because the two states

$$|\pm p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{\pm ipx/\hbar} \tag{22.7}$$

are degenerate, we can choose

$$|\pm_p\rangle = \frac{|p\rangle \pm |-p\rangle}{\sqrt{2}} \tag{22.8}$$

as energy eigenstates. These are energy and parity eigenstates, but are not momentum eigenstates.

By contrast, $[\boldsymbol{L}, \Pi] = 0$, so we can simultaneously diagonalize both orbital angular momentum and parity. Under a parity operation, the spherical angles (θ, ϕ) are sent to

$$(\theta, \phi) \to (\pi - \theta, \phi + \pi)\,. \tag{22.9}$$

This tells us that states with a definite angular position transform as

$$\Pi|\theta, \phi\rangle = |\pi - \theta, \phi + \pi\rangle\,. \tag{22.10}$$

We can write the orbital angular momentum eigenstates $|\ell, m\rangle$ in terms of these states using the matrix elements

$$\langle\theta, \phi|\ell, m\rangle := Y_{\ell,m}(\theta, \phi)\,, \tag{22.11}$$

which are known as *spherical harmonics*.

The spherical harmonic $Y_{0,0}$ is a constant, meaning that

$$\Pi|\ell = 0, m = 0\rangle = |\ell = 0, m = 0\rangle\,. \tag{22.12}$$

The $\ell = 1$ states transform together as a vector, i.e., as linear combinations of x, y, z. In particular, the $\ell = 1, m = +1$ state transforms like $x + iy$; the $\ell = 1, m = -1$ states transforms like $x - iy$; and the $\ell = 1, m = 0$ state transforms like z. Because vectors are odd under parity, this tells us that

$$\Pi|\ell = 1, m\rangle = -|\ell = 1, m\rangle\,, \tag{22.13}$$

or equivalently,

$$Y_{\ell,m}(\pi - \theta, \phi + \pi) = -Y_{\ell,m}(\theta, \phi)\,. \tag{22.14}$$

In general, $Y_{\ell,m}$ has parity $(-1)^\ell$.

22.1.4 Selection Rules

Let $\mathcal{O}$ be an operator with definite parity, i.e.

$$\Pi\mathcal{O}\Pi = \lambda\mathcal{O}\,, \tag{22.15}$$

with $\lambda = \pm 1$. Consider the matrix elements $\langle\psi|\mathcal{O}|\psi'\rangle$ of this operator with two parity eigenstates $|\psi\rangle$ and $|\psi'\rangle$, such that

$$\Pi|\psi\rangle = s|\psi\rangle\,, \quad \Pi|\psi'\rangle = s'|\psi'\rangle\,, \tag{22.16}$$

with $s, s' = \pm 1$. We then have

$$\begin{aligned}\langle\psi|\mathcal{O}|\psi'\rangle &= \langle\psi|\Pi\Pi\mathcal{O}\Pi\Pi|\psi'\rangle \\ &= \lambda ss'\langle\psi|\mathcal{O}|\psi'\rangle\,,\end{aligned} \tag{22.17}$$

where in the first step we have used $\Pi^2 = \mathbb{1}$. This implies that $\langle\psi|\mathcal{O}|\psi'\rangle = 0$ unless $\lambda ss' = 1$. Thus, if $\mathcal{O}$ is even under parity ($\lambda = +1$), then $|\psi\rangle$ and $|\psi'\rangle$ must have the same parity for the matrix element to be nonzero; similarly, if $\mathcal{O}$ is odd under parity ($\lambda = -1$), then $|\psi\rangle$ and $|\psi'\rangle$ must have opposite parity for the matrix element to be nonzero. This is a selection rule.

For example, $\langle\psi|\boldsymbol{x}|\psi'\rangle \neq 0$ only when $|\psi\rangle$ and $|\psi'\rangle$ have opposite parity. As a corollary, we see that the expectation value of $\boldsymbol{x}$ in any parity eigenstate must be zero.

22.2 Time Reversal

Classical physics is time-reversal invariant: Newton's law

$$m\ddot{\boldsymbol{x}} = -\nabla V(\boldsymbol{x}) \tag{22.18}$$

is invariant under $t \to -t, \boldsymbol{x} \to \boldsymbol{x}$. Thus, if $\boldsymbol{x}(t)$ is a valid solution to Newton's equation, then so is $\boldsymbol{x}(-t)$.

Consider now the Schrödinger equation,

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x})\right)\psi(\boldsymbol{x},t)\,. \tag{22.19}$$

If we let $t \to -t$, we see that we can get a solution if we take

$$\psi(\boldsymbol{x},t) \to \psi^*(\boldsymbol{x},-t) \tag{22.20}$$

for some solution $\psi(\boldsymbol{x},t)$. This suggests that we should take the time reversal operator Θ to be anti-unitary. (Recall that an anti-unitary operator A can be written in the form $A = KU$, where K is complex conjugation and U is some unitary operator.) Thus, we have

$$\theta(a|\alpha\rangle + b|\beta\rangle) = a^*\theta|\alpha\rangle + b^*\theta|\beta\rangle \tag{22.21}$$

for any $a, b \in \mathbb{C}$ and $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$.

We now consider combinations of time reversal and time translation operations. Assuming that time reversal is a symmetry, we require that

$$|\psi(-\delta t)\rangle = \theta|\psi(\delta t)\rangle\,, \quad |\psi(0)\rangle = \theta|\psi(0)\rangle\,. \tag{22.22}$$

By using forward and backward time translations from $t = 0$, we see that

$$\begin{aligned} |\psi(-\delta t)\rangle &= \left(1 + \frac{iH}{\hbar}\delta t\right)|\psi(0)\rangle\,, \\ |\psi(\delta t)\rangle &= \left(1 - \frac{iH}{\hbar}\delta t\right)|\psi(0)\rangle\,, \end{aligned} \tag{22.23}$$

so the statements $|\psi(-\delta t)\rangle = \theta|\psi(\delta t)\rangle$ and $|\psi(0)\rangle = \theta|\psi(0)\rangle$ imply that

$$iH\theta|\psi(0)\rangle = \theta(-iH)|\psi(0)\rangle\,. \tag{22.24}$$

Thus, we have

$$iH\theta = \theta(-iH) \tag{22.25}$$

as an operator equation. Because θ is anti-unitary, this tells us that $[H, \theta] = 0$, exactly as expected of a symmetry of the Hamiltonian.

As usual, operators transform under time reversal as $\mathcal{O} \to \theta\mathcal{O}\theta^{-1}$. An operator $\mathcal{O}$ is even/odd under time reversal if

$$\theta\mathcal{O}\theta^{-1} = \pm\mathcal{O}\,. \tag{22.26}$$

We require that

$$\theta\boldsymbol{x}\theta^{-1} = \boldsymbol{x}\,, \quad \theta\boldsymbol{p}\theta^{-1} = -\boldsymbol{p}\,. \tag{22.27}$$

There are two ways to see that $\boldsymbol{p}$ must be odd under time reversal: first, we could consider the position space representation $\boldsymbol{p} \to -i\hbar\nabla$, and use that fact that θ is anti-unitary; alternatively, we

can use the fact that time reversal should preserve the commutation algebra $[x_i, p_j] = i\hbar\delta_{ij}$, which requires that $\boldsymbol{p}$ be odd because $\boldsymbol{x}$ is even and the right-hand side contains a factor of i.

Similarly, in order to preserve the commutation algebra $[J_i, J_j] = i\hbar\epsilon_{ijk}J_k$, we need

$$\theta \boldsymbol{J} \theta^{-1} = -\boldsymbol{J}\,, \tag{22.28}$$

meaning that $\boldsymbol{L} \to -\boldsymbol{L}$ and $\boldsymbol{S} \to -\boldsymbol{S}$ under parity. Note that $\boldsymbol{L}^2 \to \boldsymbol{L}^2$, so this tells us that $\theta|\ell, m\rangle \propto |\ell, -m\rangle$. In particular,

$$\theta|\ell, m\rangle = (-1)^m|\ell, -m\rangle\,. \tag{22.29}$$

Here, the phase factor $(-1)^m$ is a convention choice built into the definition of the spherical harmonics,

$$Y^*_{\ell,m}(\theta, \phi) = (-1)^m Y_{\ell,-m}(\theta, \phi)\,. \tag{22.30}$$

22.2.1 Time Reversal and Spin

We find interesting outcomes when acting on spin-$\frac{1}{2}$ systems (or systems with other half-integer spin) with time reversal. The statement

$$\theta J_z \theta^{-1} = -J_z \tag{22.31}$$

implies for a spin-$\frac{1}{2}$ particle that

$$J_z\theta|+\rangle = -\theta J_z|+\rangle = -\frac{\hbar}{2}\theta|+\rangle\,. \tag{22.32}$$

Thus, we see that $\theta|+\rangle \propto |-\rangle$. In general, there can be some phase η, so that

$$\theta|+\rangle = \eta|-\rangle\,. \tag{22.33}$$

We can write this equation in the form

$$\theta|+\rangle = \eta e^{-i\pi S_y/\hbar}|+\rangle\,. \tag{22.34}$$

We could have similarly chosen S_x instead of S_y, or indeed any spin operator in the x, y-plane. Based on this statement, we write

$$\theta = \eta e^{-i\pi S_y/\hbar}K\,, \tag{22.35}$$

where we have included K because θ is anti-unitary. We then have

$$\theta|-\rangle = \eta e^{-i\pi S_y/\hbar}K|-\rangle = -\eta|+\rangle\,. \tag{22.36}$$

From this, we see that

$$\theta^2|+\rangle = \theta(\eta|-\rangle) = \eta^*\theta|-\rangle = -|\eta|^2|+\rangle = -|+\rangle\,, \tag{22.37}$$

where we have used the fact that $|\theta|^2 = 1$ because θ is purely a phase. Similarly, $\theta^2|-\rangle = -|-\rangle$. This means that

$$\theta^2 = -\mathbb{1} \tag{22.38}$$

holds as an operator equation for a spin-$\frac{1}{2}$ system. This is true for any system with half-integer spin. There is a standard phase choice for spin-$\frac{1}{2}$, which is to take $\eta = i$, which gives

$$\theta = ie^{-i\pi S_y/\hbar}K\,. \tag{22.39}$$

Lecture 23 (Nov. 29, 2017)

23.1 Consequences of Time Reversal Symmetry

23.1.1 Spinless Particles

Theorem 5. *Consider a system of spinless particles. Suppose that H is time-reversal invariant, and that there exists a non-degenerate energy eigenket $|n\rangle$. The corresponding energy eigenfunction can be chosen to be real.*

Proof. Because $|n\rangle$ is an energy eigenket, we have

$$H|n\rangle = E_n|n\rangle\,. \tag{23.1}$$

Because time reversal is a symmetry, we then have

$$H(\theta|n\rangle) = \theta H|n\rangle = E_n(\theta|n\rangle)\,, \tag{23.2}$$

which implies

$$|n\rangle = \eta\theta|n\rangle \tag{23.3}$$

for some phase η, because $|n\rangle$ is non-degenerate. The corresponding wavefunction is then

$$\psi_n(x) = \langle x|n\rangle = \eta\langle x|\theta|n\rangle = \eta\psi_n^*(x)\,. \tag{23.4}$$

We can then make the redefinition $\tilde{\psi}_n(x) = \eta^{-1/2}\psi_n(x)$ in order to get a real wavefunction. This completes the proof. □

23.1.2 No Conservation Law

There is no conservation law associated with time reversal. Even though $[\theta, H] = 0$, there is no concept of a time reversal quantum number. Despite the fact that θ commutes with the Hamiltonian, we find that

$$\theta U(t, t_0) \neq U(t, t_0)\theta\,, \tag{23.5}$$

where $U(t, t_0)$ is the time-evolution operator. For example, suppose that we consider an energy eigenstate,

$$H|\psi\rangle = E|\psi\rangle\,. \tag{23.6}$$

Then,

$$|\psi(t)\rangle = e^{-iEt/\hbar}|\psi(0)\rangle\,, \tag{23.7}$$

while

$$\theta|\psi(t)\rangle = e^{iEt/\hbar}|\psi(0)\rangle \neq |\psi(t)\rangle\,. \tag{23.8}$$

Thus, $\theta|\psi\rangle = |\psi\rangle$ is not preserved under time evolution.

23.1.3 Kramer's Rule for Half-Integer Spin

Theorem 6 (Kramer's Rule)**.** *Time-reversal symmetry implies a two-fold degeneracy all energy eigenstates.*

Proof. Suppose we have an energy eigenket $|n\rangle$,

$$H|n\rangle = E_n|n\rangle\,. \tag{23.9}$$

If time-reversal is a symmetry, then $\theta|n\rangle$ has the same energy E_n. Suppose that

$$\theta|n\rangle = e^{i\delta}|n\rangle\,. \tag{23.10}$$

Then,

$$\begin{aligned}\theta^2|n\rangle &= \theta\Big(e^{i\delta}|n\rangle\Big)\\ &= e^{-i\delta}\theta|n\rangle\\ &= e^{-i\delta}e^{i\delta}|n\rangle\\ &= |n\rangle\,.\end{aligned} \tag{23.11}$$

However, for half-integer spin, we have seen that $\theta^2 = -\mathbb{1}$, so this is a contradiction. Thus, $\theta|n\rangle$ must be a distinct state from $|n\rangle$. Thus, we get a two-fold degeneracy for all energy eigenstates. There may be additional degeneracy, but all states will come in pairs in this way. □

23.2 Uses of Symmetry in Solving the Schrödinger Equation

23.2.1 Symmetric Double-Well Potential

Consider some symmetric double-well potential $V(x)$ and Hamiltonian

$$H = \frac{p^2}{2m} + V(x)\,. \tag{23.12}$$

By definition, we have $V(x) = V(-x)$, because the potential is symmetric.

This Hamiltonian is invariant under both Π and θ. The eigenfunctions can thus be chosen to be real (from time-reversal invariance) and either even or odd under $x \to -x$ (from parity invariance). We can argue that the ground state $\psi_0(x)$ must be symmetric under $x \to -x$ (specifically, it has no nodes); furthermore, it is physically apparent that the ground state should be peaked near the bottom of each well. To see the first point, imagine that the ground state were antisymmetric; in this case, we must have $\psi(0) = 0$. The squared slope of the wavefunction gives the contribution to the kinetic energy. We can then cut the ground state apart at $x = 0$, negate one side, and then glue the halves back together to give a symmetric function (with a small amount of smoothing at the origin to make the new wavefunction continuously differentiable). The resulting wavefunction has the same potential energy as the supposed ground state wavefunction, but has a smaller slope at the origin, meaning it has lower energy than the ground state, which is a contradiction. Using this approach repeatedly, we can argue that the ground state has no nodes.

What can we say about the first excited state $\psi_1(x)$? Note that if we try to construct another wavefunction with zero nodes, it will not be orthogonal to the ground state. The next-lowest energy state will have the fewest nodes possible, from the argument above, and so we thus expect the first excited state to be antisymmetric with a single node.

Classically, a particle moving in such a potential will oscillate around the bottom of one of the two wells; it must choose a particular well to reside in. This is a classical example of what is called a *spontaneously broken symmetry.*

23.2.2 3D Particle in a Spherically Symmetric Potential

Consider a 3D Hamiltonian

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(r)\,. \tag{23.13}$$

The spin decouples from the orbital angular momentum, and so we will ignore spin for now. We see that $[\boldsymbol{L}, H] = 0$, because $[\boldsymbol{J}, H] = 0$ and $[\boldsymbol{S}, H] = 0$. We can then choose the energy eigenstates to be eigenstates of $\boldsymbol{L}^2$ and L_z. We then label the energy eigenstates by $|n, \ell, m\rangle$, where ℓ and m specify the eigenvalues of $\boldsymbol{L}^2$ and L_z, respectively. The wavefunction can then be written in the form

$$\psi_{n,\ell,m}(\boldsymbol{x}) = \langle \boldsymbol{x}|n, \ell, m\rangle = R_{n,\ell}(r)Y_{\ell,m}(\theta, \phi)\,. \tag{23.14}$$

That is, the radial and angular parts of the wavefunction factorize. We will see the details of this in a moment.

In polar coordinates, we can write

$$\begin{aligned} -\nabla^2 &= -\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\boldsymbol{L}^2}{r^2} \\ &= -\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{\boldsymbol{L}^2}{r^2}\,. \end{aligned} \tag{23.15}$$

Thus, the Hamiltonian acts as

$$\begin{aligned} HR_{n,\ell}(r)Y_{\ell,m}(\theta, \phi) &= \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R_{n,\ell}}{\mathrm{d}r}\right) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}R_{n,\ell} + V(r)R_{n,\ell}\right]Y_{\ell,m} \\ &= E_nR_{n,\ell}Y_{\ell,m}\,. \end{aligned} \tag{23.16}$$

We can cancel $Y_{\ell,m}$ from both sides of this equation, and it becomes an ordinary differential equation for $R_{n,\ell}(r)$.

From the requirement that $\langle n, \ell, m|n, \ell, m\rangle = 1$, we have

$$\int_0^\infty \mathrm{d}r\; r^2R^*_{n,\ell}(r)R_{n,\ell}(r)\underbrace{\int \mathrm{d}\Omega\; Y^*_{\ell,m}(\theta, \phi)Y_{\ell,m}(\theta, \phi)}_{1} = 1\,, \tag{23.17}$$

giving

$$\int_0^\infty \mathrm{d}r\; r^2R^*_{n,\ell}(r)R_{n,\ell}(r) = 1\,. \tag{23.18}$$

We can understand this ODE more easily by making the change of variables

$$R_{n,\ell}(r) = \frac{u_{n,\ell}(r)}{r}\,. \tag{23.19}$$

The radial equation then becomes

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u_{n,\ell}}{\mathrm{d}r^2} + V_{\text{eff}}(r)u_{n,\ell} = E_nu_{n,\ell}\,, \tag{23.20}$$

where

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2\ell(\ell+1)}{2mr^2} \tag{23.21}$$

and

$$\int_0^\infty \mathrm{d}r\; |u_{n,\ell}(r)|^2 = 1\,. \tag{23.22}$$

This effective potential has a local minimum (for sufficiently non-singular potentials $V(r)$), which serves to localize the energy eigenfunctions.

23.3 Approximation Methods

23.3.1 Time-Independent Perturbation Theory

Suppose that we have a Hamiltonian of the form

$$H = H_0 + V\,, \tag{23.23}$$

where H_0 is a Hamiltonian with known spectrum and V is small. We can try to find the energy eigenstates and corresponding energy eigenvalues of the full Hamiltonian by treating V as a perturbation, and writing these quantities as a systematic power series in V.

As a formal bookkeeping device, let's introduce a parameter λ, and write

$$H = H_0 + \lambda V\,. \tag{23.24}$$

We will then expand in powers of λ (ultimately, we are interested in $\lambda = 1$, so we will set it to be so at the end). Choose an orthonormal energy eigenbasis of H_0,

$$H_0|n_0\rangle = E_{n,0}|n_0\rangle\,, \quad \langle m_0|n_0\rangle = \delta_{m_0,n_0}\,. \tag{23.25}$$

We then assume we can expand the energy eigenstates of H in the form

$$|n\rangle = |n_0\rangle + \lambda|n_1\rangle + \lambda^2|n_2\rangle + \cdots\,, \tag{23.26}$$

and the corresponding energy eigenvalues in the form

$$E_n = E_{n,0} + \lambda E_{n,1} + \lambda^2 E_{n,2} + \cdots\,. \tag{23.27}$$

If we plug these expansions into the Schrödinger equation, $H|n\rangle = E_n|n\rangle$, we will get an infinite number of equations from matching the two sides order by order in λ. Explicitly, we have

$$(H_0 + \lambda V)(|n_0\rangle + \lambda|n_1\rangle + \cdots) = (E_{n,0} + \lambda E_{n,1} + \cdots)(|n_0\rangle + \lambda|n_1\rangle + \cdots)\,. \tag{23.28}$$

Lecture 24 (Dec. 4, 2017)

24.1 Non-degenerate Time-Independent Perturbation Theory

We now consider the problem of approximating the spectrum for a Hamiltonian of the form

$$H = H_0 + V\,, \tag{24.1}$$

where H_0 is a Hamiltonian with a known spectrum, and V is a "small" perturbation. This is the problem we began discussing in the last lecture. As a bookkeeping device, we introduced a parameter λ, and wrote

$$H = H_0 + \lambda V\,. \tag{24.2}$$

This parameter will allow us to organize our power counting analysis, and we can set $\lambda = 1$ at the end.

We have

$$H_0|n_0\rangle = E_{n,0}|n_0\rangle\,, \quad \langle m_0|n_0\rangle = \delta_{m_0,n_0}\,, \tag{24.3}$$

and we wish to solve

$$(H_0 + \lambda V)|n\rangle = E_n|n\rangle\,. \tag{24.4}$$

We expand the energy eigenstates of H in the form

$$|n\rangle = |n_0\rangle + \lambda|n_1\rangle + \lambda^2|n_2\rangle + \cdots\,, \tag{24.5}$$

and the corresponding energy eigenvalues in the form

$$E_n = E_{n,0} + \lambda E_{n,1} + \lambda^2 E_{n,2} + \cdots\,. \tag{24.6}$$

This gives us

$$(H_0 + \lambda V)(|n_0\rangle + \lambda|n_1\rangle + \cdots) = (E_{n,0} + \lambda E_{n,1} + \cdots)(|n_0\rangle + \lambda|n_1\rangle + \cdots)\,. \tag{24.7}$$

24.1.1 The First-Order Energy Shift

We can compare the two sides of Eq. (24.7) order-by-order in λ. At $O(\lambda^0)$, we get

$$H_0|n_0\rangle = E_{n,0}|n_0\rangle\,, \tag{24.8}$$

which we already knew. At $O(\lambda^1)$, we find

$$V|n_0\rangle + H_0|n_1\rangle = E_{n,0}|n_1\rangle + E_{n,1}|n_0\rangle\,. \tag{24.9}$$

If we take the inner product of this equation with $|n_0\rangle$, then we find

$$\langle n_0|V|n_0\rangle + \langle n_0|H_0|n_1\rangle = \langle n_0|E_{n,0}|n_1\rangle + E_{n,1}\,. \tag{24.10}$$

The second term on the left equals the first on the right, and so we are left with

$$E_{n,1} = \langle n_0|V|n_0\rangle\,. \tag{24.11}$$

This result makes sense, even classically: if we perturb the Hamiltonian, the leading-order shift in the energy of an eigenstate is just the expectation of the perturbation in the unperturbed eigenstate.

24.1.2 The First-Order Correction to the Eigenstate

Next we want to find $|n_1\rangle$, the leading-order correction to the energy eigenstate. Rearranging the $O(\lambda^1)$ terms of Eq. (24.7), we have

$$(H_0 - E_{n,0})|n_1\rangle = (E_{n,1} - V)|n_0\rangle\,. \tag{24.12}$$

If we insert

$$\mathbb{1} = \sum_{m_0} |m_0\rangle\langle m_0| \tag{24.13}$$

on both sides of this equation, the left-hand side becomes

$$(H_0 - E_{n,0})|n_1\rangle = \sum_{m_0}(H_0 - E_{n,0})|m_0\rangle\langle m_0|n_1\rangle = \sum_{m_0}(E_{m,0} - E_{n,0})|m_0\rangle\langle m_0|n_1\rangle\,, \tag{24.14}$$

while the right-hand side becomes

$$(E_{n,1} - V)|n_0\rangle = \sum_{m_0}|m_0\rangle\langle m_0|(E_{n,1} - V)|n_0\rangle = \sum_{m_0}|m_0\rangle(E_{n,1}\langle m_0|n_0\rangle - \langle m_0|V|n_0\rangle)\,. \tag{24.15}$$

We now assert that we can choose $\langle n_0|n_1\rangle = 0$; you can check that this is possible by considering the normalization of $|n\rangle$. The left-hand side then becomes

$$(H_0 - E_{n,0})|n_1\rangle = \sum_{m_0\neq n_0}(E_{m,0} - E_{n,0})|m_0\rangle\langle m_0|n_1\rangle\,, \tag{24.16}$$

while the right-hand side becomes

$$(E_{n,1} - V)|n_0\rangle = \sum_{m_0}|m_0\rangle(\langle n_0|V|n_0\rangle\delta_{n_0m_0} - \langle m_0|V|n_0\rangle) = -\sum_{m_0\neq n_0}|m_0\rangle\langle m_0|V|n_0\rangle\,. \tag{24.17}$$

Comparing these expressions, we find

$$|n_1\rangle = \sum_{m_0\neq n_0}\frac{|m_0\rangle\langle m_0|V|n_0\rangle}{E_{n,0} - E_{m,0}}\,. \tag{24.18}$$

Note that this expression only makes sense when the energy eigenvalues of H_0 are not degenerate.

We can now deduce a condition for the validity of perturbation theory: we demand that $\langle n_1|n_1\rangle \ll 1$ so that the first correction to the energy eigenstates is small, i.e., that

$$|\langle m_0|V|n_0\rangle| \ll |E_{n,0} - E_{m,0}| \tag{24.19}$$

and that their ratio decreases rapidly enough with increasing m_0.

24.1.3 The Second-Order Energy Shift

Now, we are prepared to compute $E_{n,2}$, the second-order correction to the energy. At $O(\lambda^2)$ in Eq. (24.7), we have

$$H_0|n_2\rangle + V|n_1\rangle = E_{n,0}|n_2\rangle + E_{n,1}|n_1\rangle + E_{n,2}|n_0\rangle\,. \tag{24.20}$$

If we take the inner product with $|n_0\rangle$, we can use

$$\langle n_0|H_0|n_2\rangle = E_{n,0}\langle n_0|n_2\rangle\,, \tag{24.21}$$

to eliminate the first term from each side, leaving us with

$$\langle n_0|V|n_1\rangle = E_{n,1}\cancelto{0}{\langle n_0|n_1\rangle} + E_{n,2}\langle n_0|n_0\rangle\,. \tag{24.22}$$

Because $|n_0\rangle$ is normalized, this gives us

$$E_{n,2} = \langle n_0|V|n_1\rangle = \sum_{m_0\neq n_0} \frac{|\langle n_0|V|m_0\rangle|^2}{E_{n,0}-E_{m,0}}\,. \tag{24.23}$$

Again, this only makes sense if the energy eigenstates of H_0 are non-degenerate. To second order, we then have

$$E_n = E_{n,0} + \langle n_0|V|n_0\rangle + \sum_{m_0\neq n_0} \frac{|\langle n_0|V|m_0\rangle|^2}{E_{n,0}-E_{m,0}} + O\big(V^3\big)\,. \tag{24.24}$$

24.2 Examples of Time-Independent Perturbation Theory

24.2.1 Spin in a Magnetic Field

Consider a spin-$\frac{1}{2}$ in a magnetic field of the form

$$\boldsymbol{B} = B_x\hat{\boldsymbol{x}} + B_0\hat{\boldsymbol{z}}\,, \tag{24.25}$$

with $B_x \ll B_0$ (assume that both B_x and B_0 are positive). The Hamiltonian is

$$H = -\gamma\boldsymbol{B}\cdot\boldsymbol{S} = -\gamma(B_0S^z + B_xS^z) = -\gamma\frac{\hbar}{2}(B_0\sigma^z + B_x\sigma^x)\,. \tag{24.26}$$

To analyze this system using perturbation theory, we define

$$H_0 = -\frac{\gamma\hbar}{2}B_0\sigma^z\,, \quad V = -\frac{\gamma\hbar}{2}B_x\sigma^x\,. \tag{24.27}$$

To zeroth order, the energy eigenstates are $|+\rangle$ with energy $-\gamma\hbar B_0/2$ and $|-\rangle$ with energy $+\gamma\hbar B_0/2$. To second order, the shift in the energy of the state $|+\rangle$ is

$$\begin{aligned}\Delta E_+ &= \langle +|V|+\rangle + \frac{|\langle -|V|+\rangle|^2}{E_+ - E_-}\\ &= -\frac{\gamma\hbar B_x}{2}\cancelto{0}{\langle +|\sigma^x|+\rangle} + \left(\frac{\gamma\hbar B_x}{2}\right)^2\frac{|\langle -|\sigma^x|+\rangle|^2}{-\gamma\hbar B_0}\\ &= \frac{\gamma\hbar}{2}\left(-\frac{B_x^2}{2B_0}\right).\end{aligned} \tag{24.28}$$

Similarly, we find

$$\Delta E_- = \frac{\gamma\hbar}{2}\left(\frac{B_x^2}{2B_0}\right). \tag{24.29}$$

Thus, the higher energy eigenvalue increases and the lower energy eigenvalue decreases. This is a general property of non-degenerate, time-independent perturbation theory: the second-order perturbation will always lower the energy of the ground state, because the numerator will be positive (as it is the sum of squares) and the denominator will be negative (as the ground state has lower energy than each other state).

We know the exact energy eigenvalues for this system are

$$\mp\frac{\gamma\hbar}{2}\sqrt{B_0^2+B_x^2} \approx \mp\frac{\gamma\hbar}{2}B_0\left(1+\frac{B_x^2}{2B_0^2}+\cdots\right) = \mp\frac{\gamma\hbar}{2}B_0 + \frac{\gamma\hbar}{2}\left(\mp\frac{B_x^2}{2B_0}\right) + O\big(B_x^3\big)\,, \tag{24.30}$$

which exactly matches our calculation from perturbation theory.

24.2.2 The Quadratic Stark effect

Consider an atom in a weak uniform external electric field. What is the effect on the ground state energy? We will not worry about the issue of ionization, as the likelihood of ionization from applying a small electric field is very small. Thus, we ignore the possibility that the electron can escape from the atom.

We take the unperturbed Hamiltonian to be

$$H_0 = \frac{p^2}{2m} + V_0(r)\,, \tag{24.31}$$

With V_0 rotationally symmetric, e.g.

$$V_0 = -\frac{e^2}{r} \tag{24.32}$$

for the hydrogen atom. We then take the perturbing Hamiltonian to be

$$V = -eEz\,, \tag{24.33}$$

where we have simply defined the z-axis to be the one pointing in the direction of the electric field.

The energy shift of the ground state is then

$$\Delta E_0 = \langle 0|V|0\rangle + \sum_{n\neq 0} \frac{|\langle n|V|0\rangle|^2}{E_0 - E_n} + \cdots\,. \tag{24.34}$$

The first term vanishes by symmetry arguments, using either parity or rotational symmetry. Thus, the shift starts at second order. We then have

$$\Delta E_0 = e^2E^2 \sum_{n\neq 0} \frac{|\langle n|z|0\rangle|^2}{E_0 - E_n} + \cdots\,. \tag{24.35}$$

The second-order term is negative. We define α, called the *atomic polarizability*, such that

$$\Delta E_0 = -\frac{1}{2}\alpha E^2 + \cdots\,. \tag{24.36}$$

We could have written this form down simply by symmetry arguments, but now we have a way to calculate its value:

$$\alpha = 2e^2 \sum_{n\neq 0} \frac{|\langle n|z|0\rangle|^2}{E_n - E_0}\,. \tag{24.37}$$

In order to proceed further, we need to know the particular value of $V_0(r)$ for the system we are trying to solve. For the hydrogen atom, it turns out that this calculation can be done exactly. Dimensional analysis gives us most of the answer: we have

$$\alpha \sim e^2\,\frac{\text{length}^2}{\text{energy}}\,. \tag{24.38}$$

The only length scale in the hydrogen atom is the Bohr radius a_0, and the only energy scale is e^2/a_0, which is a Rydberg. Thus, we have

$$\alpha \sim \frac{e^2a_0^2}{e^2/a_0} \sim a_0^3\,. \tag{24.39}$$

For hydrogen, the exact answer turns out to be

$$\alpha = 4.5a_0^3\,. \tag{24.40}$$

24.2.3 Van der Waals Interaction

An important interaction in chemistry and biology is the van der Waals force. Given two (hydrogen) atoms in their respective ground states, separated by a distance $r \gg a_0$ (with a_0 the Bohr radius), how does the total ground state energy depend on r?

Take the direction of separation to be the z-direction. We will take the two protons to be fixed in position. The full Hamiltonian is

$$\begin{aligned} H &= H_0 + H_1 \,, \\ H_0 &= -\frac{\hbar^2}{2m}\big(\nabla_1^2 + \nabla_2^2\big) - \frac{e^2}{r_1} - \frac{e^2}{r_2} \,, \\ H_1 &= \frac{e^2}{r} + \frac{e^2}{|\boldsymbol{r} + \boldsymbol{r}_2 - \boldsymbol{r}_1|} - \frac{e^2}{|\boldsymbol{r} + \boldsymbol{r}_2|} - \frac{e^2}{|\boldsymbol{r} - \boldsymbol{r}_1|} \,, \end{aligned} \tag{24.41}$$

where $\boldsymbol{r}$ is the vector separating the protons, and $\boldsymbol{r}_i$ are the vectors separating the electrons from their respective protons.

If $r \gg a_0$, then we can expand H_1 in powers of $\boldsymbol{r}_i/r$. This is a multipole expansion. The leading term in this case is a dipole–dipole interaction,

$$\begin{aligned} H_1 &= \frac{e^2}{r^3}(x_1x_2 + y_1y_2 - 2z_1z_2) + O\bigg(\frac{1}{r^4}\bigg) \\ &= \frac{e^2}{r^3}(\boldsymbol{r}_1 \cdot \boldsymbol{r}_2 - 3(\boldsymbol{r}_1 \cdot \hat{\boldsymbol{r}})(\boldsymbol{r}_2 \cdot \hat{\boldsymbol{r}})) + O\bigg(\frac{1}{r^4}\bigg) \,. \end{aligned} \tag{24.42}$$

We then perturb in H_1. The zeroth order ground state is

$$\psi_{\mathrm{gd}} = \psi^1_{\mathrm{gd}}(\boldsymbol{r}_1)\psi^2_{\mathrm{gd}}(\boldsymbol{r}_2) \,. \tag{24.43}$$

We then have

$$\begin{aligned} \Delta E^{(1)} &= \langle \mathrm{gd}|H_1(\boldsymbol{r}_1, \boldsymbol{r}_2)|\mathrm{gd}\rangle \\ &= e^2\langle \mathrm{gd}|\frac{x_1x_2 + y_1y_2 - 2z_1z_2}{r^3}|\mathrm{gd}\rangle \,. \end{aligned} \tag{24.44}$$

Note that

$$\langle \mathrm{gd}|x_1x_2|\mathrm{gd}\rangle = \langle (\mathrm{gd})_1|x_1|(\mathrm{gd})_1\rangle\langle (\mathrm{gd})_2|x_2|(\mathrm{gd})_2\rangle = 0 \,, \tag{24.45}$$

and similarly for all terms in $\Delta E^{(1)}$. Thus, the leading-order energy shift comes in at second order.

We find

$$\Delta E^{(2)} = \frac{e^4}{r^6}\sum_{k\neq 0}\frac{|\langle k_1, k_2|x_1x_2 + y_1y_2 - 2z_1z_2|0,0\rangle|^2}{E_{0,0} - E_{k_1,k_2}} \,. \tag{24.46}$$

Dimensional analysis tells us that

$$\Delta E^{(2)} \sim -\frac{e^4}{r^6}\frac{a_0^4}{\mathrm{Ry}} \sim -\frac{e^2a_0^5}{r^6} \,, \tag{24.47}$$

where the Rydberg is e^2/a_0. The punchline is that there is a weak attractive interaction, which is the van der Waals interaction. This is a purely quantum mechanical effect, arising because the dipole–dipole interaction induces virtual transitions to the excited states in each atom.

Lecture 25 (Dec. 6, 2017)

25.1 Degenerate Perturbation Theory

Previously, when discussing perturbation theory, we restricted ourselves to the case where the unperturbed energy levels were not degenerate. We will now consider the case where the unperturbed Hamiltonian has degeneracies. Suppose the eigenstate $|n_0\rangle$ is degenerate with $|k_0\rangle$, i.e., $E_{n,0} = E_{k,0}$. We clearly cannot use the formula from last class for the first-order correction to the eigenstate,

$$|n\rangle = |n_0\rangle + \lambda \sum_{m\neq n} \frac{|m_0\rangle\langle m_0|V|n_0\rangle}{E_{n,0} - E_{m,0}}\,, \tag{25.1}$$

because the denominator vanishes at $m = k$.

The solution is to restrict our attention to the degenerate subspace and then diagonalize the perturbation on this subspace. Within the subspace of degenerate states of the unperturbed Hamiltonian of energy $E_{n,0}$ (referred to as the *degenerate manifold of* $E_{n,0}$), we diagonalize V. Within this subspace, $H_0 \propto \mathbb{1}$ by definition (since every state in this space has the same energy eigenvalue). Thus, $PH_0P \propto \mathbb{1}$, where P is the projector onto the degenerate subspace, and so PH_0P commutes with PVP. We can then diagonalize the projected V simultaneously with the projected H_0. In the orthonormal basis that diagonalizes PVP, the off-diagonal matrix elements of PVP are zero, so we can return to using the formulas of ordinary perturbation theory. The terms where $E_{n,0} - E_{k,0}$ vanish do not occur, as the numerator also vanishes.

25.1.1 Linear Stark Effect

As an example, consider an atom in an electric field. In the case of a hydrogen atom, we found that the energy of the electron in the atom changed quadratically in the electric field (for small electric field). This was the phenomenon known as the quadratic Stark effect. Now we will discuss a situation in which the energy changes linearly in the applied electric field.

Let us choose $\boldsymbol{E} = E\hat{\boldsymbol{z}}$. Now, we consider the $n = 2$ levels of a hydrogenic atom. There are four degenerate states (ignoring spin) of the form $|n, \ell, m\rangle$, which are

$$\underbrace{|2,1,1\rangle\,,\quad |2,1,0\rangle\,,\quad |2,1,-1\rangle}_{2p}\,,\quad \underbrace{|2,0,0\rangle}_{2s}\,. \tag{25.2}$$

Our perturbation is $V = -eEz$. Consider V in the subspace of the four $n = 2$ levels. We need to compute

$$\langle 2, \ell', m'|z|2, \ell, m\rangle\,. \tag{25.3}$$

Because $[z, J_z] = 0$, we have

$$0 = \langle 2, \ell', m'|(zJ_z - J_z z)|2, \ell, m\rangle = (m - m')\langle 2, \ell', m'|z|2, \ell, m\rangle\,. \tag{25.4}$$

Thus,

$$\langle 2, \ell', m'|z|2, \ell, m\rangle = 0\,, \quad \text{if } m' \neq m\,. \tag{25.5}$$

Note also that under parity, $z \to -z$, so

$$\langle n, \ell, m'|z|n, \ell, m\rangle = 0\,, \tag{25.6}$$

i.e., we will only find a non-vanishing matrix element for $\ell' \neq \ell$.

Thus, the only non-vanishing matrix elements are

$$\langle 2,1,0|z|2,0,0\rangle = \langle 2,0,0|z|2,1,0\rangle^* \,. \tag{25.7}$$

We can compute these entries exactly, because we know the hydrogen atom wavefunctions. Explicitly, we find

$$V = \begin{pmatrix} 0 & 3ea_0E & 0 & 0 \\ 3ea_0E & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{25.8}$$

where a_0 is the Bohr radius. This matrix has eigenvalues $0, 0, \pm 3ea_0E$. Specifically, the states $|2,1,1\rangle$ and $|2,1,-1\rangle$ have no energy shift, while the state $(|2,0,0\rangle + |2,1,0\rangle)/\sqrt{2}$ has shift $+3ea_0E$ and the state $(|2,0,0\rangle - |2,1,0\rangle)/\sqrt{2}$ has shift $-3ea_0E$. This energy shift is linear in the strength of the electric field, hence the effect is called the linear Stark effect.

In general, it may be the case that diagonalizing V in the space of degenerate states does not split the degeneracy. In such a case, we must go to higher order in perturbation theory, until we are able to successfully split the degeneracy, after which point we can proceed by non-degenerate perturbation theory. If PVP does not split the degeneracy, then at the next order we will be considering operators of the form

$$PVP + \sum_{\substack{m \\ E_m \neq E_0}} P \frac{V|m\rangle\langle m|V}{E_0 - E_m} P \,. \tag{25.9}$$

25.2 Time-Dependent Perturbation Theory

We now consider time-dependent perturbations to the Hamiltonian. We begin with the Schrödinger equation,

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle \,, \tag{25.10}$$

with

$$H = H_0 + H_1(t) \,, \tag{25.11}$$

where H_0 has a known spectrum and H_1 is some weak, time-dependent perturbation.

We may want to compute the probability that the initial H_0 eigenstate $|i\rangle$ transitions to the final H_0 eigenstate $|f\rangle$ in time t. If $H_1 = 0$, then the solution to the Schrödinger equation is of the form

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_{n,0}t/\hbar}|n_0\rangle \,, \tag{25.12}$$

where the c_n are time-independent. If $H_1(t) \neq 0$, then there is additional time dependence coming from the perturbation; we can still expand any state in the basis of unperturbed energy eigenstates, so we can write the solutions to the Schrödinger equation in the form

$$|\psi(t)\rangle = \sum_n c_n(t) e^{-iE_{n,0}t/\hbar}|n_0\rangle \,, \tag{25.13}$$

where the coefficients $c_n(t)$ are now functions of time. We could absorb the time dependence in the exponential into the coefficients $c_n(t)$, but it is notationally convenient not to do so.

Applying the Schrödinger equation to Eq. (25.13),

$$(i\hbar\partial_t - H_0 - H_1)|\psi(t)\rangle = 0 \,, \tag{25.14}$$

we are left with

$$\sum_n [i\hbar\partial_t c_n(t) - H_1(t)c_n(t)]e^{-iE_{n,0}t/\hbar}|n_0\rangle = 0\,. \tag{25.15}$$

Taking the inner product with

$$\langle f_0|e^{iE_{f_0}t/\hbar}\,, \tag{25.16}$$

we arrive at

$$i\hbar\frac{\mathrm{d}c_f}{\mathrm{d}t} = \sum_n \langle f_0|H_1(t)|n_0\rangle e^{i\omega_{fn}t}c_n(t)\,, \tag{25.17}$$

where

$$\omega_{fn} = \frac{E_{f,0} - E_{n,0}}{\hbar}\,. \tag{25.18}$$

We can recognize this as a concrete example of working in the interaction picture; we have removed the time-dependence due to the unperturbed Hamiltonian, so the only time-dependence we see remaining comes from H_1.

What we have done so far is exact. We now assume that $|\psi(t = t_0)\rangle = |i\rangle$ and solve perturbatively in H_1. In the initial state,

$$c_n(t = t_0) = \delta_{ni}\,. \tag{25.19}$$

At first order, we then have

$$\frac{\mathrm{d}c_f}{\mathrm{d}t} = -\frac{i}{\hbar}\langle f_0|H_1(t)|i_0\rangle e^{i\omega_{fi}t}\,, \tag{25.20}$$

which can be integrated to yield

$$c_{fi} = \delta_{fi} - \frac{i}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\,\langle f_0|H_1(t')|i_0\rangle e^{i\omega_{fi}t'}\,. \tag{25.21}$$

This is the transition amplitude between the states $|i_0\rangle$ and $|f_0\rangle$ to first order in the perturbation H_1.

25.2.1 SHO in a Time-Dependent Electric Field

Consider a simple harmonic oscillator in a time-dependent electric field

$$\epsilon(t) = \epsilon_0 e^{-t^2/\tau^2}\,. \tag{25.22}$$

The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 - ex\epsilon(t)\,. \tag{25.23}$$

If our initial state is the ground state, $|i\rangle = |0\rangle$, then we want to compute the transition amplitude to the state $|n\rangle$ after a time t. To first order, we find this probability to be

$$P_{0\to n}(t) = \frac{1}{\hbar^2}\left|\int_{-\infty}^{t}\mathrm{d}t'\,\langle n|H_1(t')|0\rangle e^{i\omega nt'}\right|^2, \quad n \neq 0\,. \tag{25.24}$$

Here, we have used $E_{n,0} - E_{0,0} = \omega n$ for the SHO. In the limit $t \to \infty$, this becomes

$$P_{0\to n}(t\to\infty) = \frac{1}{\hbar^2}\left|\int_{-\infty}^{\infty}\mathrm{d}t'\,\langle n|H_1(t')|0\rangle e^{i\omega nt'}\right|^2, \quad n \neq 0\,. \tag{25.25}$$

We can write the perturbation in the form

$$H_1(t) = -e\epsilon(t)\sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right), \tag{25.26}$$

and so we find

$$P_{0\to n}(t\to\infty) = \frac{e^2}{2m\omega\hbar}\left|\int_{-\infty}^{\infty} \mathrm{d}t'\, \epsilon\big(t'\big)e^{i\omega n t'}\langle n|\Big(a + a^\dagger\Big)|0\rangle\right|^2, \quad n \neq 0\,. \tag{25.27}$$

Because $a|0\rangle = 0$ and $a^\dagger|0\rangle = |1\rangle$, we see that to this order, $P_{0\to n} = 0$ for $n > 1$. For $n = 1$, we find

$$\begin{aligned} P_{0\to 1}(t\to\infty) &= \frac{e^2}{2m\hbar\omega}\left|\int_{-\infty}^{\infty} \mathrm{d}t'\; e^{i\omega t'}\epsilon_0 e^{-t^2/\tau^2}\right|^2 \\ &= \frac{\pi e^2\epsilon_0^2}{2m\hbar\omega}\tau^2 e^{-\omega^2\tau^2/4}\,. \end{aligned} \tag{25.28}$$

If we work to higher orders, we will find nonzero transitions to higher excitation states $|n\rangle$. Note that $P_{0\to 1}$ is small both for $\omega\tau \gg 1$ (slow pulse) and $\omega\tau \ll 1$ (fast pulse), and peaks when $\omega\tau \sim 1$.

25.2.2 Second-Order Transition Amplitude

Iterating the solution for $c_n(t)$ to next order, we find

$$c_n^{(2)}(t) = \left(-\frac{i}{\hbar}\right)^2 \sum_m \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\; e^{i\omega_{nm}t'} H_{1,nm}(t') e^{i\omega_{mi}t''} H_{1,mi}(t'')\,, \tag{25.29}$$

where

$$H_{1,jk}(t) = \langle j|H_1(t)|k\rangle\,. \tag{25.30}$$

Incorporating the higher-order terms, the transition probability is

$$P_{i\to n} = \left|c_n^{(1)}(t) + c_n^{(2)}(t) + \cdots\right|^2. \tag{25.31}$$

Lecture 26 (Dec. 11, 2017)

26.1 Harmonic Perturbations

We now study perturbations of the form

$$H_1(t) = Ve^{-i\omega t} + V^\dagger e^{i\omega t}\,. \tag{26.1}$$

Here, V is some operator acting on the Hilbert space that depends on the degrees of freedom in the system. We assume that V is "weak," so that we can hope to treat it perturbatively.

If this perturbation acts for a long time, we expect that it will induce a transition from the initially prepared state $|i\rangle$ to some final state $|f\rangle$ where the system has absorbed energy $\hbar\omega$ from the perturbation, i.e.,

$$E_f - E_i = \hbar\omega\,. \tag{26.2}$$

In the last class, we computed the transition amplitude for an arbitrary time-dependent perturbation at first order:

$$c_{fi}(t) = -\frac{i}{\hbar}\int_0^t \mathrm{d}t'\ \langle f|H_1|i\rangle e^{i\omega_{fi}t'}\,. \tag{26.3}$$

For the case of harmonic perturbations, we find

$$\begin{aligned} c_{fi}(t) &= -\frac{i}{\hbar}\int_0^t \mathrm{d}t'\ \left[\langle f|V|i\rangle e^{i(\omega_{fi}-\omega)t'} + \langle f|V^\dagger|i\rangle e^{i(\omega_{fi}+\omega)t'}\right] \\ &= \frac{1}{\hbar}\left[\langle f|V|i\rangle\left(\frac{1-e^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right) + \langle f|V^\dagger|i\rangle\left(\frac{1-e^{i(\omega_{fi}+\omega)t}}{\omega_{fi}+\omega}\right)\right]. \end{aligned} \tag{26.4}$$

As $t\to\infty$, $|c_{fi}|^2$ is appreciable if either $\omega_{fi}-\omega\approx 0$, i.e., $E_f\approx E_i+\hbar\omega$ (absorption), or if $\omega_{fi}+\omega\approx 0$, i.e., $E_f = E_i - \hbar\omega$ (emission).

We now specialize to the case where ω is tuned so that

$$E_f - E_i \approx \hbar\omega\,. \tag{26.5}$$

In this case, the second term in Eq. (26.4) is small compared to the first, so we can write

$$c_{fi}(t) \approx \frac{1}{\hbar}\langle f|V|i\rangle\left(\frac{1-e^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right). \tag{26.6}$$

The transition probability is then

$$\begin{aligned} P_{fi}(t) &= |c_{fi}(t)|^2 \\ &= \frac{1}{\hbar^2}|\langle f|V|i\rangle|^2\frac{\left|1-e^{i(\omega_{fi}-\omega)t}\right|^2}{(\omega_{fi}-\omega)^2} \\ &= \frac{1}{\hbar^2}|\langle f|V|i\rangle|^2\frac{\sin^2\left(\frac{(\omega_{fi}-\omega)t}{2}\right)}{\left(\frac{(\omega_{fi}-\omega)t}{2}\right)^2}t^2\,. \end{aligned} \tag{26.7}$$

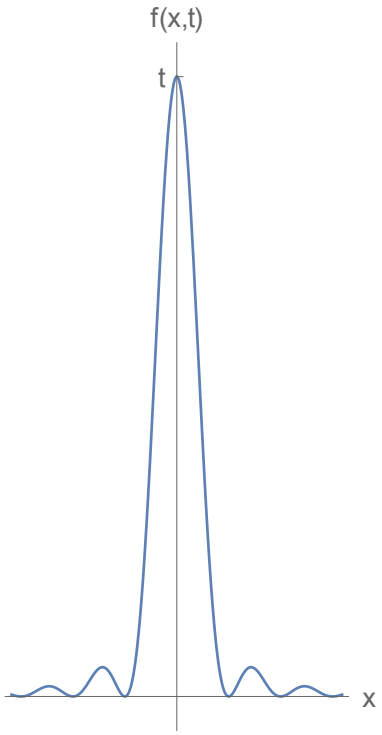
In this last line, we have simply used

$$\begin{aligned} \left|1-e^{i\theta}\right|^2 &= \left|e^{i\theta/2}\left(e^{-i\theta/2}-e^{i\theta/2}\right)\right|^2 \\ &= \left|-2i\sin\frac{\theta}{2}\right|^2 \\ &= 4\sin^2\frac{\theta}{2}\,. \end{aligned} \tag{26.8}$$

Now, consider the function

$$f(x,t) = \frac{\sin^2 tx}{tx^2}\,, \tag{26.9}$$

which is shown below:



At $x = 0$, we have $f(0,t) = t$. As $x \to \infty$, $f(x,t) \to 0$. The function oscillates with a period proportional to $1/t$. As $t \to \infty$, the function decays to zero very quickly, but $f(0,t) = t \to \infty$. Thus, this appears to be approaching something like a delta function. We can check this by noting that

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}x\, \frac{\sin^2 tx}{tx^2} &= \int_{-\infty}^{\infty} \frac{\mathrm{d}(tx)}{t}\, \frac{\sin^2 tx}{tx^2} \\ &= \int_{-\infty}^{\infty} \mathrm{d}u\, \frac{\sin^2 u}{u^2} \\ &= \pi\,,\end{aligned} \tag{26.10}$$

which we see is independent of t. We conclude that

$$\lim_{t\to\infty} \frac{\sin^2 tx}{tx^2} = \pi\delta(x)\,. \tag{26.11}$$

Thus,

$$\lim_{t\to\infty} P_{fi}(t) = \frac{1}{\hbar^2}|\langle f|V|i\rangle|^2 \pi\delta\left(\frac{\omega_{fi}-\omega}{2}\right)t = \frac{2\pi}{\hbar^2}|\langle f|V|i\rangle|^2\delta(\omega_{fi}-\omega)t\,. \tag{26.12}$$

This may seem odd, because the result we found is that as $t \to \infty$, the probability of transition diverges to infinity, because it is proportional to t. The interpretation here is that there is a constant transition rate,

$$R_{fi} = \frac{P_{fi}}{t} = \frac{2\pi}{\hbar^2}|\langle f|V|i\rangle|^2\delta(\omega_{fi}-\omega)\,. \tag{26.13}$$

In general, there will be several possible final states $|f\rangle$ to which H_1 can induce a transition. The total transition rate is then given by the sum over all final states $|f\rangle$ of this rate. We define the density of states $\rho(E)$ such that $\rho(E)\,\mathrm{d}E$ is the number of states between E and $E+\mathrm{d}E$. Then the net transition rate out of the initial state is given by

$$\begin{aligned} W_{i,\text{out}} &= \sum_f R_{fi} \\ &= \int \mathrm{d}E_f\ \rho(E_f)\frac{2\pi}{\hbar^2}|\langle f|V|i\rangle|^2\delta(\omega_f-\omega_i-\omega)\,. \end{aligned} \tag{26.14}$$

Changing variables in the delta function, we reach

$$W_{i,\text{out}} = \int \mathrm{d}E_f\ \rho(E_f)\frac{2\pi}{\hbar}|\langle f|V|i\rangle|^2\delta(E_f-E_i-\hbar\omega)\,. \tag{26.15}$$

This famous result is known as *Fermi's Golden Rule*.

26.1.1 The Photoelectric Effect

Consider an electron in a hydrogen atom interacting with an external EM field. This system has Hamiltonian

$$H = \frac{\left(\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}\right)^2}{2m} - \frac{e^2}{r}\,, \tag{26.16}$$

where $\boldsymbol{A}$ is the external vector potential. Take the vector potential to be of the form

$$\boldsymbol{A} = \boldsymbol{A}_0\cos(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})\,, \tag{26.17}$$

and assume that $\boldsymbol{A}_0$ is weak. We can then write

$$H = \underbrace{\frac{\boldsymbol{p}^2}{2m} - \frac{e^2}{r}}_{H_0} - \frac{e}{2mc}(\boldsymbol{p}\cdot\boldsymbol{A}+\boldsymbol{A}\cdot\boldsymbol{p}) + \frac{e^2}{2mc^2}\boldsymbol{A}^2\,. \tag{26.18}$$

where H_0 is the hydrogen atom Hamiltonian. The final term, proportional to $\boldsymbol{A}^2$, does not contribute to the problem we are considering at first order in perturbation theory, so we will simply drop it here.

It is convenient to work in Coulomb gauge, $\nabla\cdot\boldsymbol{A}=0$. In this gauge,

$$(\boldsymbol{p}\cdot\boldsymbol{A}+\boldsymbol{A}\cdot\boldsymbol{p})\psi = (-i\hbar\nabla\cdot\boldsymbol{A}+2\boldsymbol{A}\cdot\boldsymbol{p}) = 2\boldsymbol{A}\cdot\boldsymbol{p}\psi\,. \tag{26.19}$$

Thus, we can use

$$H = H_0 + H_1 \tag{26.20}$$

with

$$\begin{aligned} H_1 &= -\frac{e}{mc}\cos(\omega t-\boldsymbol{k}\cdot\boldsymbol{r})\boldsymbol{A}_0\cdot\boldsymbol{p} \\ &= -\frac{e}{2mc}\left(e^{i(\omega t-\boldsymbol{k}\cdot\boldsymbol{r})}+e^{-i(\omega t-\boldsymbol{k}\cdot\boldsymbol{r})}\right)\boldsymbol{A}_0\cdot\boldsymbol{p}\,. \end{aligned} \tag{26.21}$$

The first term gives the transition rate proportional to $\delta(E_f-E_i+\hbar\omega)$, while the second term gives the transition rate proportional to $\delta(E_f-E_i-\hbar\omega)$.

If we take the atom to initially be in its ground state, then only the second term can contribute, so Fermi's Golden Rule gives

$$R_{0\to f} = \frac{2\pi}{\hbar}\sum_{\boldsymbol{k}_f}\left|\langle \boldsymbol{k}_f|\frac{e}{2mc}\boldsymbol{A}_0\cdot\boldsymbol{p}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}|0\rangle\right|^2\delta(E_f-E_0-\hbar\omega)\,. \tag{26.22}$$

Here, $|0\rangle$ denotes the ground state of the hydrogen atom, with wavefunction

$$\psi_0(\boldsymbol{x}) = \langle \boldsymbol{x}|0\rangle = \frac{1}{\left(\pi a_0^3\right)^{1/2}}e^{-r/a_0}\,, \tag{26.23}$$

where a_0 is the Bohr radius. The ket $|\boldsymbol{k}_f\rangle$ is a plane wave state with wave vector $\boldsymbol{k}_f$, which describes the electron that is kicked out of the atom by the perturbation.

We can compute the matrix element in the transition rate by putting the system in a box of side length L, which gives

$$\begin{aligned}\frac{e}{2mc}\boldsymbol{A}_0\cdot\int \mathrm{d}^3x\,\frac{e^{-i\boldsymbol{k}_f\cdot\boldsymbol{r}}}{L^{3/2}}(-i\hbar\nabla)\frac{e^{-r/a_0}}{\left(\pi a_0^3\right)^{1/2}} &= \frac{e}{2mc}\boldsymbol{A}_0\cdot\int\frac{\mathrm{d}^3x}{L^{3/2}}\left(i\hbar\nabla\left(e^{-i\boldsymbol{k}_f\cdot\boldsymbol{r}}\right)\right)\frac{e^{-r/a_0}}{\left(\pi a_0^3\right)^{1/2}}\\ &= \frac{e\hbar}{2mc}\frac{\boldsymbol{A}_0\cdot\boldsymbol{k}_f}{L^{3/2}\left(\pi a_0^3\right)^{1/2}}\int\mathrm{d}^3x\,e^{-i\boldsymbol{k}_f\cdot\boldsymbol{r}}e^{-r/a_0}\\ &= \frac{e\hbar}{2mc}\frac{\boldsymbol{A}_0\cdot\boldsymbol{k}_f}{L^{3/2}\left(\pi a_0^3\right)^{1/2}}\frac{8\pi}{a_0\left(k_f^2+1/a_0^2\right)^2}\,.\end{aligned} \tag{26.24}$$

Note that we have ignored the factor of $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ in this matrix element; this was not an accident. For typical wavelengths involved in photo-ionization, $|\boldsymbol{k}|\ll|\boldsymbol{k}_f|$.

For the sum over final states, we make the replacement

$$\sum_{\boldsymbol{k}_f}\to L^3\int\frac{\mathrm{d}^3k_f}{(2\pi)^3}\,. \tag{26.25}$$

The logic is that, in a box of size L, the set of allowed momenta has spacing $2\pi/L$, and so each differential element in the integral is of the form

$$\frac{\mathrm{d}k_x}{(2\pi/L)}\,, \tag{26.26}$$

which gives the form of the replacement above.

Putting everything together, we then have

$$\begin{aligned}R_{0\to f} &= \frac{2\pi}{\hbar}\frac{L^3}{L^3}\int\frac{\mathrm{d}^3k_f}{(2\pi)^3}\left(\frac{e\hbar}{2mc}\right)^2\frac{(\boldsymbol{A}_0\cdot\boldsymbol{k}_f)^2}{\pi a_0^3}\frac{(8\pi)^2}{a_0^2\left(k_f^2+1/a_0^2\right)^4}\delta(E_f-E_0-\hbar\omega)\\ &= \frac{e^2\hbar}{2m^2c^2a_0^5}\int\frac{\mathrm{d}k_f\;k_f^2\,\mathrm{d}\Omega}{(2\pi)^3}\frac{(\boldsymbol{A}_0\cdot\boldsymbol{k}_f)^2(8\pi)^2}{\left(k_f^2+1/a_0^2\right)^4}\delta\left(\frac{\hbar^2k_f^2}{2m}-E_0-\hbar\omega\right)\\ &= \frac{e^2k_f}{2mc^2a_0^5\hbar}\int\frac{\mathrm{d}\Omega}{(2\pi)^3}\frac{(\boldsymbol{A}_0\cdot\boldsymbol{k}_f)^2(8\pi)^2}{\left(k_f^2+1/a_0^2\right)^4}\,,\end{aligned} \tag{26.27}$$

with k_f determined by energy conservation. The only angular dependence is in $\boldsymbol{A}_0 \cdot \boldsymbol{k}_f$, so using

$$\int \mathrm{d}\Omega \cos^2\theta = \frac{4\pi}{3} \tag{26.28}$$

we reach the final result,

$$R_{0\to f} = \frac{16e^2 k_f^3 A_0^2}{3mc^2\hbar a_0^5} \frac{1}{\left(k_f^2 + 1/a_0^2\right)^4} \,. \tag{26.29}$$